

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 23:04:17 ; Search time 4272.77 Seconds

(without alignments)
17041.707 Million cell updates/sec

Title: US-10-027-000-3

Perfect score: 2502

Sequence: 1 atgctgatgatgatgttga.....attgtgtctgcgcgtgtaa 2502

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	427	17.1	4193	8	KFBGLUC
2	406.2	16.2	38269	1	AF486581
3	393.6	15.7	39305	8	SPU31009
4	393.6	15.7	41799	8	SPBC1683
5	393.6	15.7	85837	8	SPU33010
6	392.2	15.7	2793	8	AF329731
7	350.4	14.0	3578	1	AF005277
8	343.4	13.7	3710	1	ATUC861
9	302.2	12.1	2635	8	AB1293760
10	270	10.8	110000	1	REU80928.2
11	205	8.2	2601	6	AX489445
12	197.4	7.9	37445	1	SC8A6
13	184.4	7.4	3241	6	AR173332
14	184.4	7.4	17665	1	AF521878
15	178.2	7.1	2760	1	SEY14327
16	175.2	7.0	12441	1	AF079762
17	161	6.4	2256	6	E11369
18	157.2	6.3	2640	1	TNGBL
19	156.8	6.3	3179	1	CEGDGB
20	152.2	6.1	11763	1	AE005774
21	149.4	6.0	14878	1	AE005649
22	143.6	5.7	10615	1	AE012240
23	138.6	5.5	15052	1	AF053579
24	138.6	5.5	50937	6	AR159871
25	137.4	5.5	21243	1	AE001690
26	136.2	5.4	3265	1	RAU92808
27	135.8	5.3	2166	6	AR205082
28	132.6	5.3	3158	1	RALBGLT
29	129.2	5.2	2075	1	ANGEMA
30	127.6	5.1	2620	1	CSBGLT
31	127.6	5.1	4537	1	AF090429
32	126.4	5.1	2346	1	PA1276021
33	126	5.0	12140	1	AE011777
34	123.2	4.9	110000	2	LMFICR36_08
35	118.2	4.7	3542	1	AB071166
36	117.2	4.7	3955	1	MIBBGA
37	114.4	4.6	11557	1	AE004598
38	111.6	4.5	15852	1	AE006929
39	111.6	4.5	36300	1	MFC128
40	109	4.4	13332	1	AE005788
41	106.2	4.2	2370	1	PRO35425
42	104.4	4.2	2894	1	CTBGLB
43	104.4	4.2	12882	1	AE013715
44	104.4	4.2	150754	9	AC023491
45	104.4	4.2	258050	1	AJ414153

ALIGNMENTS

RESULT 1

KFBGLUC 4193 bp DNA linear PLN 12-SEP-1993

LOCUS KFBGLUC

DEFINITION Kluveromyces fragilis beta-glucosidase gene.

ACCESSION X05918

VERSION X05918.1 GI:2804

KEYWORDS beta-glucosidase.

SOURCE Kluveromyces marxianus.

ORGANISM Kluveromyces marxianus

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Kluveromyces.

REFERENCE 1 (bases 1 to 4193)

AUTHORS Raynal,A., Gerbaud,C., Francinques,M.C. and Guerlneau,M.

TITLE Sequence and transcription of the beta-glucosidase gene of Kluveromyces fragilis cloned in Saccharomyces cerevisiae

Pred. No. is the number of results predicted by chance to have a

JOURNAL MEDLINE PUBMED	Cur. Genet. 12 (3), 175-184 (1987)
88210533	
2835179	
FEATURES	
source	Location/Qualifiers
	1. .4193
	/organism="Kluyveromyces marxianus"
	/strain="ATCC 12424"
	/db_xref="taxon:4911"
	/clone_lib="pHCG3"
	268. .276
repeat_region	/note="direct repeat 1"
	307. .315
repeat_region	/note="direct repeat 1"
	316. .320
promoter	/note="TATA-Box like sequence"
	342. .355
promoter	/note="TATA-Box like sequence"
	400. .408
repeat_region	/note="direct repeat 2"
	406. .412
promoter	/note="TATA-Box like sequence"
	416. .424
repeat_region	/note="direct repeat 2"
	459
misc_feature	/note="pot. transcription start site"
	482. .483
misc_feature	/note="pot. alternate transcription start site"
	612. .3149
CDS	/note="beta-glucosidase (AA 1 - 845)"
	/codon_start=1
	/protein_id="CAA29353.1"
	/db_xref="GI:2805"
	/translation="MSKFDVEQLISELNODEKISLSAVDFWHTKKIERGIPARVYS DGPNGIRGTKEFDPSGCEPENGSTFEDRLLEFAGKIMAKESIAKNAAYILGPT TNMROGGGGEFSESDPEPLAGMATSVYKGMGGIAATVAKHPQNDLEODRFS NISVSEAPLREIYEPPIRYAKVHANPVCIMYAKKXNGDHSQSKLLIDIREMKK DGLMSDFRGTYTTAAAIKNGLDIEFPGPTRMRIRALVSHSREOITTEDVDYR QVLKMFVVDNLEKGTIVENGPESTSNMTERSDILRELAOSI VLLKKNNTLTSSK ERROYHVGPNAAKAKTSGGGSASMGSYVYVPEYGVNKGKEDVTGAVSHKSTGVS GLASSIIDLAKPADARNAGLIAFYENPBERSEDEPFHVTKNNSVLIPEFKHEHE KVDPKRYEFTLVLTGOVYPOBDGDYITSLQYSGGLFYINDELITDKNBERSPCE GAGKERETKTLTKKGQVYNNRVYRGSGSPGSIAGFGGAGGFOGAVKAPKAIIDDEIINN AAELAAKHDKRAVLITGLANGEMETEGYBRENADLPKRTNELRAYLAKNPVTIYNQSG TPVEFPMLEENNALVOAMYGNGNELGNAIADVLGDVVPNGELTSLWPFKLDQNAFLIN FKTEFGRVYGEDIFVGYRYEYKLRKRVAFPGYGLSYTFEELDISDKYTDKIDISL VDVNTGDKFAGESEVOVYFSAIINSKVSYPVKELGFEKVELBEGEKTVNIELEKDK AIVSENEIKGMHNEAGBYLTVSGTSSDDILSYKEFVKNOLYMKGL"
repeat_unit	3314. .3322
	/note="inverted repeat A"
misc_feature	3442. .3445
	/note="yeast termination consensus sequence"
misc_feature	3472. .3474
	/note="yeast termination consensus sequence"
misc_feature	3498. .3501
	/note="yeast termination consensus sequence"
misc_feature	3541. .3543
	/note="yeast termination consensus sequence"
repeat_unit	3642. .3652
	/note="inverted repeat B"
repeat_unit	3655. .3666
	/note="inverted repeat B"
repeat_unit	3714. .3722
	/note="inverted repeat A"
BASE COUNT	1338 a 675 c 941 g 1239 t
ORIGIN	
Query Match	17.1%; Score 427; DB 8; Length 4193;
Best Local Similarity	50.7%; Pred. No. 1.5e-42;
Matches 1253; Conservative	0; Mismatches 1155; Indels 63; Gaps 7;
1 ATGCTGATATTGATGTTGAGCCATCTTGTGAAGAAGCTCACCCGCGAGAGTGCAT 60	

OY	1123	TACACCGCGCGCGCTT	-CACACCGCTCTCCCACTTCAGCGAGCAGTCCCTCAGCG	1180
Db	1752	TACACCGAGGGGCTTACACAAATCGATTGGAGGTTTGGCAGAGTGGATTGATC	1811	
OY	1181	CCGACGGCGCTCCGGGATCGCGCTGGAGGGTCTTCAACA	-GCCCGCTGGTACCCCTPA	1238
Db	1812	GATCGTGCAAAACGACGAGATGCTGAAAATGCTGGATTATTTGCCAAGTTTACTGCCAT	1871	
OY	1239	CCGCCAGCACA	-----TTGACGAGCTCTTTCACCAAGACG	1275
Db	1872	CCAGTAGAAGAGATCTGGAAGATGAAGAACATTCCTCCAGTTACCAAATGCTAATAGATCC	1931	
OY	1276	GACATGCACCGTGGGACCTACTACACCCCAAGCGCGCAGACAG	-----TGGTAC	1328
Db	1932	AATGTCATCTTATTTGATTTTCAACATGAGAAAGTGGATTCAAAGAACCTTACTTTT	1991	
OY	1337	GCCGACATGAGGGCGACGTACACCGCGAGAGATGTGCACCTACAGCTGGCGCTCGTC	1386	
Db	1992	GTAACCTTAACCGGACAGTACGTGCTCCAGAAAGATGGTGATTATATCTTCAAGTTTCA	2051	
OY	1387	GTCGCGGCACGCAAAAGGCGTACGAGACGACAGCTGCTGTCACAAACGCCACAG	1446	
Db	2052	GTTATGCTCTGCTGTGTTCTACTTAAACGATGAGTGTATTGATTCACAAAGACAAAC	2111	
OY	1447	CAGTCCCCCGGCATGGCTTCTTGCGCTCGCGCACCCCGGAGAGACGGCGCATCAT	1506	
Db	2112	CAGAAGAGGGTAGTTTCTCTTTGGAGCTGTACCAAGAAAGAACCAAAAGTTGACT	2171	
OY	1507	CTCGTCAAGGGCAACAGCTACAGTTCAGATGAGATTTGGCTCGCACCCACCTACAC	1566	
Db	2172	TTGAAAGAAAGGCCCAAGTTTATATGTAAGATTGATACGTTCTTGCCCAACTTCAGGT	2231	
OY	1567	CTCAAAGGCGACACATCTGTCGCCGCGACGCTCCCTCGCGTGGCGGCTGCAAGGTC	1626	
Db	2232	TTGGTTGGGGA-----ATTGCGTGCAGTGGATGATTCAGAGTGTCTATTAAAGCG	2282	
OY	1627	ATTGACGACCGCGCGAAATCGMAAAGTCCGTGCCCTCGCCAAAGAGACGACCAAGTTC	1686	
Db	2283	ATCGATGATGACGAGGAGATTAGAAACGACGAAATTGACGCTAAGAGATATTAAGCT	2342	
OY	1687	ATCATCTGCGGGGCTTAAACGCCGATGGGAGACCGGAGGGCGCGACCGCGCAGCATG	1746	
Db	2343	GTTGTTGATTAATGGATTAAATGGTGAATGGGAACCGAAGTTTATGACGAGAAACATG	2402	
OY	1747	AAGCTCCCCGCGGTGCGAGCAGCTATTTGCCAGTGGCGCGCGCGAATCCCAACAC	1806	
Db	2403	GATTTGCCAAAAGAACAAATGAAATTAAGTTCGTGTTTGAAGCAAAATCCAAATACT	2462	
OY	1807	GTCGTCTCATGACAGACGGGACCCCCGAGAGATGCCCTGCTGAGCGCACGCCGCC	1866	
Db	2463	GTTATCGTTAACCATATGAGGTACCCCAAGTTGATGATTCCTTGGTTAGAAAGGCAATACG	2522	
OY	1867	GTCATCCAGGCGCTGTCGCGGCGCAAGAGGGGCAATCATTTCCGAGCTGCTCTT	1926	
Db	2523	CTAGTTCAAGTTGGTAGCGGTGATTAAGATTGGGTAATGCTATCGCAAGATCTTTGAC	2582	
OY	1927	GCGCATACAAACCTCTGGGCAAGCTGTCCCTCAGTTTCCCAACGCGCTGGAGAGCAAC	1986	
Db	2583	GGTGACGTGGTCCAAATGGTAAGTTATGCTCTCTTGCCCATTTAAGTGGAAATAT	2642	
OY	1987	CCGCGCTTTCACATCTTCGCGACCGGAGCGCGCGCACGCTTACGGCGAGAGCTGTAC	2046	
Db	2643	CCACGCTTTTAAACTTCAAGAACCGAGTTTCGAAAGAGTTGTTTACGATGAGATATCTT	2702	
OY	2047	GTCGGGTACAGTACTACGAGTTTGGCGACAGAGCTCAATTTCCCTTTGGCACGCG	2106	
Db	2703	GTTGGTTTATAGTACTACGAAAGCTTCAAGAAAGATGACCTTCCCTCGGATATGTT	2762	
OY	2107	CTGTCTACACACACTTTTGGCTTTTCCAAATCTCTCGGTCTTCACAAAGACGCGCAAGCTG	2166	
Db	2763	CTATCGTATACACATTCGCACTAGTATTTTTCAGCTTAAAGTAACTGATGATTAAGTA	2822	

QY	2167	ACCCTGTCCTCCCTTCGGTGAAGAACACCGGCTCCGTCGCCGCGCACAGGTGGCCACGCTC	2226
DB	2823	GATATTTCAGTTGATGTGAAGATACTGTGTATTAATTTTCGTCGCGAGTGGTGC	2882
QY	2227	TACGTCAAGCCCTCCAAAGCGGCAAGATTAAACGCCCGCTCAAGAGCTCAAGGCTTC	2286
DB	2883	GTCCTACTTCAGCGCTCTAAACTCTAAGGCTCTCGAACCGGTTAAAGAGTTC	2942
QY	2287	GCAAGGTCGACACTGCACCGCGGACAGACGAGCGGTGACATTCGAGAGGACGAGAAAG	2346
DB	2943	GAAAGAGTTCATTGGAACACGAGTGAAGAAAGACAGTTAATATTGAACCTGAAATTTGAA	3002
QY	2347	TACGTCCGTGCGCTATTTTGATGAGGAGCGGATCAGTGTGTGCGAAAGGCTGACTAT	2406
DB	3003	GATGCAATTTCTACTATTACGAAAGAGCTCGTAAATGGCAGCTGGAAGCAGGTAAATAC	3062
QY	2407	GAGTTATTCGT 2417	
DB	3063	TTGGTTTCAGT 3073	
RESULT 2			
LOCUS	AF486581	38269 bp	DNA linear BCT 21-MAR-2002
DEFINITION	uncultured bacterium aminoglycoside 3'-N-acetyltransferase gene, complete cds.		
ACCESSION	AF486581		
VERSION	AF486581.1	GI:19569565	
KEYWORDS			
SOURCE	uncultured bacterium;		
ORGANISM	uncultured bacterium		
REFERENCE	Bacteria; environmental samples.		
AUTHORS	1 (bases 1 to 38269)		
	Cappellano, C. M., Courtice, S., Ball, M., Francon, F.-X., Normand, P., Heylinck, G., Guerin, M., Martinez, A., Hopke, J., Kolvek, S., August, P. R., Osburne, M. S., Jeannin, P., Simone, P. and Pernodet, J. -L.		
TITLE	Recombinant Environmental Libraries provide Access to Microbial diversity for Natural Products Drug Discovery		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 38269)		
AUTHORS	August, P. R. and Martinez, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-FEB-2002) Molecular Biodiversity, Aventis Cambridge Genomics Center, 26 Landsdowne Street, Cambridge, MA 02139, USA		
FEATURES	Location/Qualifiers		
Source	1..38269		
	/organism="uncultured bacterium"		
	/db_xref="taxon:77133"		
	/note="Isolated from soil"		
	complement(30829..31617)		
	/note="confers resistance to kanamycin in Escherichia coli; similar to known cognate proteins"		
	/codon_start=1		
	/transl_table=1		
	/product="aminoglycoside 3'-N-acetyltransferase"		
	/protein_id="AA192107.1"		
	/db_xref="GI:19569566"		
	/translation="MRSVRAKRSLSADLSAIGLADDDVLYRALQVQKTYGPPRAITIDLRVIGAGTITIGTCOMQDELKDDPSMKPHIATDPEKRSRTDRNGTPEALRTTPGALRSPGASMAALGCGEALFADHALDGYDQSPGLKLVAEKVLMLGAPLDWTLLHHAHLEADLPENKRLIRYEAFLVDGEEVWVWFEEFSEPPDLIPEDYFATIVAEFLATGKRGKEVEASSVLPAAAMVAFGVMDLERWGKTL"		
BASE COUNT	6318 a	12117 c	12998 g 6835 t 1 others
ORIGIN			
Query Match	16.2%	Score 406.2;	DB 1; Length 38269;
Best Local Similarity	50.8%;	Pred. No. 2,7e-40;	
Matches 1213;	Conservative 0;	Mismatches 1098;	Indels 78; Gaps 7;
QY	17	TTGAGGCATCTTGAAGAAGCTACACCTCGGCCGAGAAGTGCATCTGCTGGCTGGTATACG	76

OY	77	ACTTCTGCGACACAAAGGCTCTCTCCCAAGATGGAAGTCCCTCTCTCCGTTTACAGATG	136
Dp	32304	ATTTCCTGTCGCTGCGCGGGAGATCGATGCGCTGGCGTGGCGCAAGCTCTGCGGTACCGAGC	32363
OY	137	GCCCCAAGCGCGTAAGAGGAGACCAAGTCTTCTCAATGCG---GTCCCTGCGGACCTGCTTCC	193
Dp	32364	GCGCGAAGCGGGCGCTGCGCGCGGCGCTCGCTGATCGCGGGGTCAAGTCCGCGGCGCTTCC	32423
OY	194	CTTGGCGGACGTCGCTCGGTTCCACATTCAACCAAATCTGCTCGAAGAGGCAAGGTAGA	253
Dp	32424	CGGTGGGAGATCGCGCTCGCGCCACCTGGAACGTCCGCGTCTTACGGAAGTCTGGACCG	32483
OY	254	TGATGGCCAAAGAGCGCATCTCGTAAGATGTCGCAATGATCTCCGCGCCACTATCAACA	313
Dp	32484	CTTTGGCCGAGAGAGGTGAATGAAAGGGGGCGCATGTGAGCTCTCGCGCGACGCTAACCC	32543
OY	314	TGCAAGCTTCCCTCTCGGTGAGCGTGGCTTCGAGTCGATTGTGAGAGTCCGTTCTGG	373
Dp	32544	TCCATCGTTTCGCTACCAACAGGGGCCAATTTTCGAGTGTATTTCGAGAGTACCCGATCTCA	32603
OY	374	CGGGCTTGGGAGCTGGGGCTCTCATTCGCCGGCATTCAGAGACACTGAGATGTCAGCGTACGA	433
Dp	32604	CCGCGACCTCGCGGTGCGCTATATCGAGGGGTTCGAGGGGCAAGGGCGTCTCGGCGACGA	32663
OY	434	TCAAGCACTTTTGTTCATATGATCATGAGAGAGGAGGACGACGATGATGTCAGACGATGTGA	493
Dp	32664	TCAAGCACTTTTGGCGGCATGATCCGAGATTCGAGCGAACCAACCATCAGCTCCGAGATG	32723
OY	494	CGGAGCGGGCTCTCGGTGAATCTACGCACTCCGTTCCAGATTGCTGTGCGAGACTCCC	553
Dp	32724	ACGAGCGGGCGCTGCGTGAAGCTCTACCTCGTGGCGTTGAGTGGCGGTGTAACCGGGCGG	32783
OY	554	AGCGGGGTGCTTCATGAGAGCGGGTACCAATGACATCAATGCGGTCTCTGAGCGAGAACCC	613
Dp	32784	GCACCTGGGGCATCATGAGACTCTTACAAACAAGCTCAATGAGCACTTACCGAGGAGAACCC	32843
OY	614	CTAAATATCTTGATGGGATGCTTCCAAAGGAATGGGGTTGGGATGGCTTAATATAGCG	673
Dp	32844	CTGTGGTCTCGACCGAGTGTGCGCAAGGATGGCACTATGAGCGGTGATGTCTCG	32903
OY	674	ACTGTGAGGACATACATGACTACCAACAGAAAGCGTTGTGACAGGGCTCGAGCTTGAAGATCC	733
Dp	32904	ACTGTGTCGGCTCGCACTGCACTGCGCGCGGACCGCTCAATGACCGGGCTCGACTTGAATATCC	32963
OY	734	CCGAGCTCCACGCTTCGAGGAGAAACCTCAAGTTCAAGCTCTCCACGGAAGCCCT	793
Dp	32964	CCGGGCGCGCGCGCACCTGCGCGGGAAGCTGTGGCGCGGTGAGGCGCGCGAGCTCA	33023
OY	794	TTATTCACGTCATTGACACAGAGGGCTAAGGAAAGTTCTTCAAGTTGTCGAAGAAATGTGTCG	853
Dp	33024	GCGCGGAACGCTGCGGAGCCCGGGGTGCTCGAGGTGCTCGGGCTATTGAGCGGACCGGG	33083
OY	854	CTTCCGAGTGAAGAGAAAGCGCCCGGAGAGACTGTCAACAAACCCCGGAAACGGCAG	913
Dp	33084	CGCTGAGCG---ATCAACAGGAGTGGGTGAAGCGGGCGCGAGTCCGCGGAGCACCGGG	33140
OY	914	CTCTCTCTCGGAAGTTGGCAACAGAGGCACTGTCCTGTGAAGACGACAAACAGTTC	973
Dp	33141	CGCTGATCTCTCGCGCGCGGCGGGAAGGGCGGCTCTCTGTAAAAAGTCGGGGGAAGTACG	33200
OY	974	TGCCCTTGACCAAGAAAGAAAGACGCTGATGTGTGGGCCCAACGCCAACACAGGCCACAT	1033
Dp	33201	TGCAATTGACCGGGTGGCGCAAGATCGGGGTGATCGGGCCGAAACGCTTAAAGTGGCGACA	33260
OY	1034	ACCAAGCGAGGAGCTCTGCGGACTCAGGGCCCTATACAGCAGTCACTCTTTCAGGCC	1093
Dp	33261	TGATGGCGCGGACACGCGGACGAGCTCAACCCGCACTATCGGGTTTCGCTGGGATGTCTC	33320
OY	1094	TGACGACGAGCTCGAGAGCCGCCCATGTATACAGCTGGGGCTTACACAGCCGTTCTCTC	1153
Dp	33321	T-----GTTGCGGGCTTGGGCGAAGAGCGCTTACGCTATGACGG	33358

Db 34416 CGAAGTGTGACAGTCTACGTCTGCGGCGGATGCG---ACGGTTCGCGGCGGCGCA 34472
 QY 2270 AGGACCTACAGGCTTCCAGAGGTCAGACGCGGAGAGAGAGGCGGTGACAA 2329
 Db 34473 AGGACCTACAGGCTTCCAGAGGTCAGACGCGGAGAGAGAGGCGGTGACAA 34532
 QY 2330 TCGAGGACGAGAGAGAGTACGTCTGCGTATTTTGAATGAGAGCGGGA 2378
 Db 34533 TGGAGCTGATGACCGGCTTTCGCTTACTACCGGACCGAGCGGACGA 34581

RESULT 3
 SPU33009 39305 bp DNA linear PLN 14-NOV-1995
 LOCUS Schizosaccharomyces pombe cosmid 1683.
 DEFINITION U33009
 ACCESSION U33009
 VERSION U33009.1 GI:1063616
 KEYWORDS
 SOURCE Schizosaccharomyces pombe.
 ORGANISM Schizosaccharomyces pombe
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomyces.
 1 (bases 1 to 39305)
 MCombie, W.R.
 REFERENCE
 TITLE Sequence analysis of a region of the fission yeast genome
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 39305)
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-1995) W. Richard MCombie, Cold Spring Harbor
 Lab, 1 Bungtown Rd, P.O. Box 100, Cold Spring Harbor, NY 11754, USA
 FEATURES
 SOURCE
 1. 39305
 /organism="Schizosaccharomyces pombe"
 /db_xref="taxon:4896"
 /chromosome="2"
 /clone="cosmid 1683"
 /clone_1bp="Mizukami, et al., 1993. Cell 73:121-132."
 BASE COUNT 12234 a 7052 c 7011 g 13008 t
 ORIGIN

Query Match 15.7%; Score 393.6; DB 8; Length 39305;
 Best Local Similarity 49.0%; Pred. No. 8.8e-39;
 Matches 1199; Conservative 0; Mismatches 1229; Indels 18; Gaps 5;

QY 1 ATGGCTGATATTGATGTTGAGGCACTTGAAGAAGCTCAGCCGCGAGAGAGTGAT 60
 Db 8497 ATGATGGAACATGATGTTGAAGATTGATCAATCAATTGATTAAGTGAAGAAGCTATG 8536
 QY 61 CTGCTGGGTGATGACTCTGACACACAAGAAGCTCTCCCAAGCATGAGTCCCTCT 120
 Db 8557 CTGCTTCGCGACTGACCTTGGCATATGACACCAATTCCTGCTTAACATCTTCC 8616
 QY 121 CTCGCTTTACAGATGCGCCCAAGCGGTAAAGAGGACCAAGTCTTCAATGGGCTCCT 180
 Db 8617 ATTAGATTATACAGAGCGGCTTAATGATCCGTTGAAGTACCTTCTTCAATTTCTACCA 8676
 QY 181 GCGGCTGCTCTCTTGGCGGACGCTGCTGCTTCCACATTCACCAAACTCTGCTCGAA 240
 Db 8677 TCCCTGTTTTCCTCTGGGACAGCGCTAGGGGCTACTTTCGACAAAAGTTACTATTC 8736
 QY 241 GAGCAGATGATGATGAGGCAAGAGGCAATGCTAAGGTGGCGATGATCTCTGCG 300
 Db 8737 GAAATGCTGAATATTATAGCAGAAAGACAAACGAAAGGTGTTAGTGTGGTTTGGGT 8796
 QY 301 CCGACTATCAACATGCAACGCTCCCTCTGCTGAGAGCTGAGTCTGATGATGATGAG 360
 Db 8797 CCAACGGTAATATCCATCGATCGAGACCTTAATGTAGGGGTTTGAATCATTTTCTGAA 8856
 QY 361 GATCCGTTCCCTGGGCGCTTGGAGCTGCGGCTCTCATCCGCGGCAATTCAGACACTGA 420

Db 8857 GATTTACTTTATTCAGGCTCTTGCAGCTAGTTANGTTATTTCTGGATTCGAAATATAAAC 8916
 QY 421 GTGAGGCTACGATTCAGACCTTTTGTGCAATGATCAGGAGGACAGGCGCATGATGAG 480
 Db 8917 GTCCAGCATGATTTAGCACTTTTGTGTGAATGATGAGAAATGAAAGAACTCTGTG 8976
 QY 481 CAGAGCATGCTCAACGAGGCGCTCTCCGAAATCTACGACATCCCTCCAGATTGCT 540
 Db 8977 AGATTGACGCTCTCACAAAGGCGACCTTAGAGAGTATTTCTTATGCAATTCAGTTGGCA 9036
 QY 541 GTGCGAGACTCCACAGCGGCTGCTTCATGACGCGCTACAGTGCATCATGCGCTGCG 600
 Db 9037 TGCAAATATTCATTAATTCATATCATCTTATCTTACAAACAGGTGATGCTGAAC 9096
 QY 601 TGCAGGAGAGACCTTAATATCTTGTGATGAGATCTCTGAAAGAAAGGGGTGGATGAG 660
 Db 9097 GTTTCCCATCTCGAATTTTATTTATGACATATTTTAAAGAAAGAGTGGAGTGAAGGT 9156
 QY 661 CTATCATGAGCAGCTGTTACGCGACATACAGTACCAAGAACCCCTTGTGCGGCGCTC 720
 Db 9157 ACTATATATCTGATTTGTTTGAACCTTATTCGTTGAAGAAACCTTATGATGAGGTTTG 9216
 QY 721 GACCTGAGATGCGCGGACCTCCAGCTTCGAGAG---AAACACTCAAGTTCAACATC 777
 Db 9217 GATTTGGAGATGCGCAAGAAAGCCGAGATTCGAAATGTCACACTTTCACACTTAGTA 9276
 QY 778 TCGACGGAAGCCCTTTATCCAGCTATGACCAAGGCTGGAAGCTTCTTACGTTTC 837
 Db 9277 GGATCCAAAGGCTTTGCGAGTCAATTTTATGAGAAAGCAAAAGATGTCGGAAGCTT 9336
 QY 838 GTCAAGAGTGTCTCTCCCTCCGAGTACGAGAGAGAGGCGCGGAGCACTGTCAACAAC 897
 Db 9337 GTAAACATCTCCGCGAGATACAGAAAGCAAAATCATTTGTAGCTAATATATGATTCG 9396
 QY 898 ACCCCGAAAGGCGAGCTCTCTCCGAAAGTTGGCAACGAGGCGATGCTGCTGAAG 957
 Db 9397 TCTTTGTTTAAAGAGAGC---ACTTAAATAATTTCAAGCAGACGCTATGCTCTTGAAA 9453
 QY 958 AAGCAACAGACGTTCTGCTCTGACAGAAAGAAAGAAACGCTGATGTCGGCCCCAAC 1017
 Db 9454 AATTAAGAAAGAACTTTTGCCATTTGCAAGAGAGGCGACCTTTCAGATTTAGACCTAAC 9513
 QY 1018 GCCAAGCAGGCGCATACAGAGGCGGAGGCTCTGCGCACTCAGGCGCTACTACGACAT 1077
 Db 9514 GCAAAGTTTGTATATACAGCGGAGGTGATCGGCAAAATTTGAAGCTTATATATGATG 9573
 QY 1078 ACTCCTTTGACGCGCTCAGCAAGACGCTGAGACGCGCCCATCTGACACCGTGGCGCC 1137
 Db 9574 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9633
 QY 1138 TACAGC-----ACGGTCTCTCCATTTCTAAGGAGACATGCTCAGCGCCAGCGCGCT 1191
 Db 9634 CATTAATCTTAACCTTCCAAATATTGCAAAATTTATGATCAACCTTAGAACCGGCAAG 9693
 QY 1192 CCGGATGCGCTGAGGCTTTCACAGAGCCCTGATACCCCTTAAACCGCAGACATTT 1251
 Db 9694 CATGATGATGCTGCAAGATTTATCTTGAACCTGCAACCTCGAAATATGAACTTTGATA 9753
 QY 1252 GACGAGCTTCTTCAACCAAGAGGACATGACCTGTGTGACTACTACACCTCCAA--G 1308
 Db 9754 GACGATTAGATTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 9813
 QY 1309 GCGGAGACAGCGGTGATGCGGACATGAGGAGGAGGAGCTACACCGCGAGAGAGTGCAC 1368
 Db 9814 AAGGATGCTATTTCTATATGACATTTGAAGTTATTTATTTCTGACAGAGAGCGTGT 9873
 QY 1369 TACGAGCTGCGCTCTGCTCTGCGGACAGCAAAAGGCTAGATGAGAGAGCTGCTGTC 1428
 Db 9874 TATGATTTGGGATATCTGTGTGGAAGCGCTGCTGTTATGATGATGATGATGATGATG 9933
 QY 1429 GTGCAACAGCAGCAAGAGGTCCTGCGGATGCTTTTGGCTCGCGCCACCGCGAG 1488
 Db 9934 ATGACACAGACAGCAAGAACCCCTACGAATCATATTTTGAATTTGAAACAAATAGAA 9993

1489 GAGAGGGGGCCGATCATCTGCTCAAGGGCAACAGTACAGTTCAGATGCTGGCC 1548
 9994 GAGAGAAATTTCTATTATTAAGAGGGGAGAAAGTAAATGATCGGGCTGAGATGGA 10053
 1549 TCCGACACCCTTACACCTTCAAGGGGACACACATGTCGCCGGCCAGCTCCCTCCG 1608
 10054 AATGCGACCTTACACCTGTCACACTATCTGCTCCAGTACGGGAGAAATGACTCT 10113
 1609 GTGCGGGCTGCAAGGCTATGACACACAGCCGAATGAAAGTCCGTGCGCTCCG 1668
 10114 ATTGCTGTGTTAAGTATGATCTCAAGAGATGATTAACGAGTGGGTTGCTCC 10173
 1669 AAGGAGCAGCAGGATGATCATCTGCGGGGCTTAACGGCGATGAGAGCCGAGGGC 1728
 10174 AATCATGTTGATGTTGATTTATGTTGTTGTTGACAGCAGATGGGAAACGGAAGT 10233
 1729 GCGGACCGGCGAGCATGAACTCCCGGCGCTGCTGACAGAGGACCCCGAGAGATGCTG 1788
 10234 GAAGATAGAAAACATGACTTTACATGCTATCTGACAAACCTGTTATTTCTATCTG 10293
 1789 GCGGGAACCCAAACAGCTGCTGCTGATGACAGAGGACCCCGAGAGATGCTG 1848
 10294 CATCCAAATCCCAATACAGTACTGTCATCAGCAGCAGCAGCAGCAGCAGCAG 10353
 1849 CTCGACCGCAGCCGCGGCTGATCCAGGCTGCTGACAGGCGGACAGCAGCAGCAGC 1908
 10354 ATCTCGAAGCCACACCTTATGACATTTGTCACAAAGATGAGATTAAGAAATGCT 10413
 1909 ATGCGGAGCTGCTGTTTGGCAGCTACAAACCCCTCGGGAGAGCTGCTGCTGCTCC 1968
 10414 TTAGCAAAATATTTTGGAGAGCAAAACCTGTTGAGAAATGCTCAATATATCTCC 10473
 1969 AAGCGCTCAGAGACACCCCGCTTCTCACTTCGACCGAGCGGCGGCGACAGCTG 2028
 10474 AAAAACTAAGATATCTCTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 10533
 2029 TACGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088
 10534 TATGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10593
 2089 TTCCCTTGGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
 10594 TTTCATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10653
 2149 CACAAGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208
 10654 AATGCTGTTGAAGGTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10713
 2209 GCACAGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2268
 10714 GCAGAAATATATCAATATATATTTCTCAATC---GTCGTTCAATGATGACCTGCT 10770
 2269 AAGGAGCTAAGGCTTCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
 10771 AAGAATCTAAAGAGTTTCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 10830
 2329 ATGAG 2388
 10831 ATGAG 10890
 2389 GTGGAAG 2434
 10891 TCTGAAG 10936

RESULT 4
 SPC1683 41799 bp DNA Linear PLN 23-AUG-2000
 LOCUS Schizosaccharomyces pombe cosmid c1683.
 DEFINITION AL355920
 ACCESSION GI:7801293
 VERSION
 KEYWORDS adenosine deaminase; alpha-glucosidase; beta-glucosidase precursor

(EC 3.2.1.21); ferric reductase transmembrane component; frp1;
 gnt4; glucose transporter protein; glycosyl hydrolase family 3;
 inosine-uridine preferring nucleoside hydrolase; isocitrate lyase
 (EC 4.1.3.1); MFS allantoate permease; MFS inorganic phosphate
 transporter; MFS transporter of unknown specificity; PCS1 allantoate
 transporter; TFI-type LTR; transcriptional activator; zinc finger.
 fission yeast.
 Schizosaccharomyces pombe
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomyces.
 1 (bases 1 to 41799)
 McCombie, W. R.
 Sequence analysis of a region of the fission yeast genome
 2 (bases 1 to 41799)
 McCombie, W. R. and Lyne, M. H.
 Direct Submission
 Submitted (11-APR-2000) W. Richard McCombie, Cold Spring Harbor
 Lab, 1 Bungtown Rd, P.O. Box 100, Cold Spring Harbor, NY 11754, USA
 and Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA
 Notes:
 Details of yeast sequencing at the Sanger Centre are available on
 the World Wide Web.
 (URL, http://www.sanger.ac.uk/projects/S_pombe/)
 CDS are numbered using the following system eg SPC5H10.01c. SP (S.
 pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
 (complementary strand).
 The length in codons is given for each CDS.
 IMPORTANT: This sequence MAY NOT be the entire insert of the
 sequenced clone. It may be shorter because we only sequence
 overlapping sections once, or longer, because we arrange for a
 small overlap between neighbouring submissions.
 Cosmid SPC1683 is overlapped at its 5' end by cosmid SPC359 and
 at its 3' end by cosmid SPC1198.
 Location/Qualifiers
 1. 41799
 /organism="Schizosaccharomyces pombe"
 /db_xref="taxon:4896"
 /chromosome="II"
 /map="f11"
 /clone="cosmids 359 (partial), 1198 and 1683"
 /clone_11b="Mizukami, et al., 1993. Cell 73:121-132."
 1. 500
 /note="nominal overlap with cosmid SPC359 S. pombe
 chromosome II"
 1351..1543
 /note="mRNA from spc08587 1 193"
 3226..4947
 /gene="SPC1683.01"
 3226..4947
 /gene="SPC1683.01"
 /note="SPC1683.01, len:573,
 SIMILARITY:Schizosaccharomyces pombe, YAC_SCHPO, putative
 inorganic phosphate transporter c2333.12., (559 aa), fasta
 scores: opt: 2395, E():0, (70.0% identity in 574 aa)
 /codon_start=1
 /label="SPC1683.01
 /product="MFS inorganic phosphate transporter"
 /protein_id="CAB91163.1"
 /translation="MKINIFKSHGDSNTAEERPVEQVDAEDQEHENFWGLTAKE
 FLMLAGVFFLDSDLEIFINLVPIFEFLVGLGKPNKGHYPSIRGLVNAS
 NININGOLMEFGMDPEGRKFEYKEMIIIVATILLAMPKSHSPSKMMVYCW
 RWLLGICGDYPMASAIRLQFGLALYPAIGVLPRLMEETQFKNSQOLNSGNDPK
 HPLNRBGRKRLKGVRLQFGLALYPAIGVLPRLMEETQFKNSQOLNSGNDPK
 TSNFEDDELKKNPSVTKGPEIHESSENTLSNTVENDEPTEKQFESVAPARS
 GFIOYFQVGHFHLGTSVCFELDIARYGVNLQNSVILKINLFGSSGNEYRTTKN
 AIGNLIITAVGVYGVYFVLEILGRKWILOQGVILGLMEALLAGMNEISTGR
 PACFVIAOLFSENFSPNSTEPIYPAVEPRAVGTAGISALGKCAIILASLEFNL
 STIGGVNMTIFCGCMGGILFTLLLPETKGRADRIDVELFYGGDGVCEKSKKS
 WYVNGIF"

```

misc_feature
join(3364..4004,4134..4872)
/gene="SPBC1683.01"
/Note="Match to PF00083 sugar_tr, Sugar (and other)
transporter Score 132.74"
complement(5308..5478)
/Note="partial degraded TFI-type LTR"
5967..6986
/gene="SPBC1683.02"
5967..6986
/gene="SPBC1683.02"
/Note="SPBC1683.02, len:339, SIMILARITY: Saccharomyces
cerevisiae, ADA_YEAST, probable adenosine deaminase, (347
aa), fasta scores: opt: 723, E():0, (37.6% identity in 351
aa)"
/codon_start=1
/label="SPBC1683.02
/product="putative adenosine deaminase"
/db_xref="GI:7801295"
/translation="MTDIERFIEKLPKALHLHLEGTLEAEELKIKLSHRNKIPKROSS
IEIKESVNEHDLASFLVEYEGVELLHEDPFYDLCYOLAKRASAQVWVYAEWEPD
QJHTRGISTEYVIGKLRARDADRDHITISQIMCTIREKSFENAEETINASIPK
SEIIGIDSNENNNPPIKFKVQARQOLGRLCHDLOKNTTHIRALDEIGV
ERIDGINILDDPELIKALERNIPFTYCPSPNEIVYQKAPERIRIMIDYKVTIN
SDPVMHCFYITENNLAKGASLTKKELVQICRNSPEAMISSEKRNHYLEALNSF
ASAYD"
6141..6956
/gene="SPBC1683.02"
/Note="Match to PF00962 A_deaminase, Adenosine/AMP
deaminase Score 164.32"
complement(7448..9007)
/gene="SPBC1683.03c"
complement(7448..9007)
/gene="SPBC1683.03c"
/Note="SPBC1683.03c, len:519, SIMILARITY: Saccharomyces
cerevisiae, Q08902, orf yoi378w., (515 aa), fasta scores:
opt: 2052, E():0, (58.0% identity in 509 aa)"
/codon_start=1
/label="SPBC1683.03c
/product="MFS transporter of unknown specificity"
/protein_id="CAB91165.1"
/db_xref="GI:7801296"
/translation="MNNYKSLGTYNTRNMEKSKDLVONEDPPLPYOKTSMLEFHE
IYEVFLTAQLOMTOAGLGOSIAPNNIIGSLGTPNPGOLSWFPASYSLVGTFLIA
GLGDIYGHKKMVEGYIMFCIMSLISGSYAKVIMEDVCRALGTGAPVILPNAI
ALGRVYPRGKKNLIRALFGATPANGSLGSVEGFRALSSMPEWTWATYVIVF
AITGTRALPHITADEVEREKQFDYGAFFVSGVLINFSMNOGQPVYVQPYVITL
IIGFLSLVFLVLEKRVYVOPILAPSMNSMESCVLICVAAGACFGIMVYILMOFLN
LRFPAPLVTATQLEPGISGCAALTTGTLRLKPRKIMVSMIAFVGTILLATAP
IHQTYWAOYFVSIIIVTWMQMDSPAAATLMLSDPYKQHOGIAASLVSTVAVSISIG
LGAGTVEHLLNHHKGVLDLRGRSAMVGTGSGGVCVALITVAFASVLYKGGQKSPK
PFLIMKNTKAY"
11335..13833
/gene="SPBC1683.04"
/Note="SPBC1683.04"
11335..13833
/gene="SPBC1683.04"
/Note="SPBC1683.04, len:832, SIMILARITY: Kluyveromyces
marxianus, BGIS_KLUMA, beta-glucosidase precursor, (845
aa), fasta scores: opt: 1274, E():0, (46.3% identity in
848 aa)"
/codon_start=1
/label="SPBC1683.04
/product="beta-glucosidase precursor (EC 3.2.1.21);
glycosyl hydrolase family 3"
/protein_id="CAB91166.1"
/db_xref="GI:7801297"
/translation="MMEHVEDLILNQLDISERKAMLLSGTDLMTHTAIPRLNIPRILS
DPSNGIRGTSFNSPSCAFPCGTALGATFDKLLPEVGEVLAEBKAKGVSVLGGPT
VINHRGLNGRPFESSEDSTSLGLAASVILIGLSKNVQACIKHFDVNDMDENSV
SIDVGBRALREVIYMPOLACKYSNPKSLMTSYKNGVGHVSQSRILDNILRKEMW
KGTISDMWCTYSIKKAIADAGIDLEMPGCPREPRNNTIOHLVGSKEISLSTDEAKN
VTLKXHSNONTBAENHCELNDSCLREALKKTSOSIVLAKNKKLLPLSRKTPA
VIGPAKVCNYSGGGSANLKPYYTVMYDGIKAIKAGPEVALGCHNLINPLNANL

```

```

misc_feature
11428..13776
/gene="SPBC1683.04"
/Note="Match to PF00933 glyco_hydro_3, Glycosyl hydrolase
family 3 Score 687.70"
16074..16191
/Note="mRNA from spc06472 118 1"
17673..19352
/gene="SPBC1683.05"
17673..19352
/gene="SPBC1683.05"
/Note="SPBC1683.05, len:559, SIMILARITY: Saccharomyces
cerevisiae, THY_YEAST, putative thiamine transporter
yoi192c., (559 aa), fasta scores: opt: 828, E():0, (26.8%
identity in 512 aa)"
/codon_start=1
/label="SPBC1683.05
/product="NCS1 allantoin transporter"
/protein_id="CAB91167.1"
/db_xref="GI:7801298"
/translation="MEDPKSDEKATDITISERNLIDVGGESSVDVYKGFASPIKTL
ELSSGPERENIDLPPTPPDRHNSALDITITLMSNGISASAFRGTSYMEKSGSPKA
LAALAGNVFLAMPTLNGLESHYHLPVAVQASVGFYFNTLILRLTAGLFTYG
TNYVGAECVOTLIYATFKSFRSKYKLPADAGITSPFLISFYVWVISPFLIRE
YLOREFLISISTYIACFAMLIPLCNVGSVWVDQPAVSGRSMVFPICALNSVA
GFTSLAVNVDFTYRVKHPKTPYVOMLLEPVAVASAPIDISGVASKINGRAMMP
LOIANWNSGRGAAPFMGITVYSLYMANIONISYVTAANDIVLEPFRILDRBAOV
IYITIGMAIYPAKTIIONGRFLAFJLSISITFLPAGTICPAKRFHHYDIDEPN
PSGILYRKLGLNLRALIAFCACVPLIPGAMINSITPMDGVILHYITGYTSPM
TAFILYKMLNVEPFAKETLEAVYPPKSNALVDPSTLSGDKFWYIDY"
complement(20737..21669)
/gene="SPBC1683.06c"
complement(20737..21669)
/gene="SPBC1683.06c"
/Note="SPBC1683.06c, len:310, SIMILARITY: Escherichia
coli., YBEK_ECOLI, hypothetical 33.8 kd protein in
leuB-gltI intergenic region., (311 aa), fasta scores: opt:
838, E():0, (44.5% identity in 308 aa)"
/codon_start=1
/label="SPBC1683.06c
/product="putative inosine-uridine preferring nucleoside
hydrolase"
Query Match
15.7%; Score 393.6; DB 8; Length 41799;
Best local similarity 49.0%; Pred. No. 8.6e-39;
Matches 1199; Conservative 0; Mismatches 1222; Indels 18; Gaps 5;

```


LOCUS	SPU33010	85837	bp	DNA	linear	PLN 14-NOV-1995
DEFINITION	Schizosaccharomyces pombe cosmid	359,	1198	and	1683.	
ACCESSION	U33010					
VERSION	U33010.1	GI:1063617				
KEYWORDS						
SOURCE						
ORGANISM	Schizosaccharomyces pombe.					
REFERENCE	Eukaryota, Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomyces.					
AUTHORS	1 (bases 1 to 85837)					
TITLE	McCombie, W.R.					
JOURNAL	Sequence analysis of a region of the fission yeast genome					
REFERENCE	Unpublished					
AUTHORS	2 (bases 1 to 85837)					
TITLE	McCombie, W.R.					
JOURNAL	Direct Submission					
REFERENCE	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor Lab, 1 Bungtown Rd, P.O. Box 100, Cold Spring Harbor, NY 11754, USA					
FEATURES	Location/Qualifiers					
SOURCE	1..85837					
BASE COUNT	27560 a 14781 c 15395 g 28101 t					
ORIGIN						
Query Match	15.7%;	Score 393.6;	DB 8;	Length 85837;		
Best Local Similarity	49.0%;	Pred. No. 7.2e-39;				
Matches 1199;	Conservative	0;	Mismatches 1229;	Indels 18;	Gaps 5;	
1	ATGCGTGAATGATGATTGAGGCGCATCTTGAAGAAGCTCACCCCTGGCGAAGAAGCTCAT	60				
29116	ATGATGGAACAAGATGTTGGAATTTGATCAATTCATTTGATATAGTATAGTAAAGCGGTATG	29175				
61	CTGCTGCGTGTATGCGACTTCTGGACACAAAGGCTCTCCCAAGCATGAGATCCCTCT	120				
29176	CTGCTTCCGGACAGCACTTTGGCATCTGAGCAATTCCTCGCTTAACATTCCTCTCC	29235				
121	CTCCGCTTTACAGATGCGCCCAAGCGGTAAGAGGAGCAAGTTCTTCAATGCGCTCCT	180				
29236	ATTAGATTATCAGAGCGCCCTAATGTAATCCCTGGAACTAGCTTCTTCAATCTTCACCA	29295				
181	GGGGCTGCTCCCTGGGCGCAAGCGCTCGGCTCCACATTTCAACCAACTCGCTCGAA	240				
29296	TCCGCTGTTTCCCTGCTGAGCAGCGCTAGGCGCTACTTCTTGACAAAAGTTACTATTTC	29355				
241	GAGGCAAGTATGATGATGGGCAAGAGGCGCATCGCTAAGATGCGCATGTATCTCTGCGC	300				
29356	GAAATTTGGTGAATATTATACAGAGAAGCAAAAGCAAGAGTATAGTGTGTGTTGGGT	29415				
301	CCGACTATTAACATGCAACGCTCCCTCTCGGCTGAGCGTGGCTTGCAGTCGATGGTAG	360				
29416	CCAAAGCGTAAATATCCATGCTGAGACCTTAAATGTAAGGAGTTTGAATCATTTCTGAA	29475				
361	GATCCGTTTCCCTGGGCGGCTTGGAGCGTGGCGCTCATCCGCGGCAATTCAGAGCACTGGA	420				
29476	GATTTACTTTATACAGCTTTTGCAGCTTAATTTATGTTATTTCTTGAGATGCAAAAGTAAAC	29535				
421	GTGCAAGCTACATCAAGCAACTTTTGTGCAATGATCAGAGAGACAGCGCGCATGATGTTG	480				
29536	GTCCAAAGCTATGATTAAGCATTTTGTGTATGATTAAGTAAAGGAACCTCTGTC	29595				
481	CAGAGCATGCTACAGAGGCGGCTCCGCTCGTAATTTAGCACTCCGCTTCAGATTGCT	540				
29596	AGTATTGACGCTCACAAAGGCGCACTTACAGAAAGTATATCTTATGTCATTTCAGTTGGCA	29655				
541	GTGCGAGACTCCAGCCGGGCTGCTTCATGACGCGCTACAAATGCAATGCGCGTGTGC	600				

QY 1669 AAGGACGACGACGATCATCTATGCGGCGCTTAACGCCGACTGGAGACCGAGGCC 1728
 DB 30793 AATCAGTTGATTTGTTGATTTTATGTTGAGGCTTGAACAGCAGATGGAAACGGAAGCT 30852
 QY 1729 GCGGACCGGCGAGCATGAGCTCCCGGCGTGGACGACCTATGCCGACTGGCC 1788
 DB 30853 GAGATAGAAAAACATGACTTTACCACCTGTTATCTGACAACTGCTTATTTATCTG 30912
 QY 1789 GCGGGAACCCAAACACCTGCTGCTATGACAGAGCGGACCCCGAGAGATGCCCTG 1848
 DB 30913 CATCCAAATCCAAATAGATGTTGACTGATGACGACATCCCAATTTAGATGCCATGG 30972
 QY 1849 CTCGACGCGACGCGCGCTCATTCAGGCTGTTAGCGGCGACAGAGCGGCAACTCC 1908
 DB 30973 ATCTGGAAGCCACACCTTATGCACTTTGGTCAACGCGTACGAAATTTAGAAATGCT 31032
 QY 1909 ATTGCGAGCTGCTTTGGGAGCTACAAACCCCGGCGAGCTCCCTCAGCTTCCCC 1968
 DB 31033 TTAGCAAACTATTTTGGAGAGCAAAACCTTGTGAAATTTGCCAATTAATTCCT 31092
 QY 1969 AAGCCCTGACAGACACCCCGCTTCTCAACTTCCGACCGAGCGGCGGACGCTG 2028
 DB 31093 AAAAACTAAGATATCTGCTCTATCTTTCATTTCTTCAAGCCGCTGCTGCTT 31152
 QY 2029 TACGCGAGAGCTCTACGCTGGGTACAGTACTAGAGTTTCCGACAGAGCTCAAT 2088
 DB 31153 TATGAGAGAGAGCTCTTGTGGTATGATATGAAAGCGTATGAAAGGAGGCTTGG 31212
 QY 2089 TTCCCTTTGGGCGGCTGCTGCTACACCTTTGCTTCCAACTCTCCGCTGCT 2148
 DB 31213 TTTCATTTGGATGATGGGCTTTCATATCTTCTTGAATTTCAATTTTATTTTGAAG 31272
 QY 2149 CACAAGAGCGGACGACGCTGCTCCCTCCGTAAGAAACACCGGCTCCGCGCCG 2208
 DB 31273 AATTGGTGGAAGATTAAGATTCATTTGGAGATTTGGAAATTCGAGCCCATGCTGCT 31332
 QY 2209 GACAGGTGGCCGCTCTACGTCAGCCCTCCAGGCGCCAGATTAACCGCCGCTC 2268
 DB 31333 GAGAAATTAATCAAGTATTAATTTCTAATATC--GTTCTGCTGATTAATTCGACTGTT 31389
 QY 2269 AAGGAGCTCAAGGCTTGGCAAGGTGCAAGCTGACGACCGGCGAGAGAGCGGCTGCA 2328
 DB 31390 AAGAGACTTAAGAGATTTTCTTAAGTGTGTTTATGCTCTTAAGAGACCAATTAATTA 31449
 QY 2339 ATCGAGAGCAGAGAGATGCTGCTGCTGATTTTATGAGAGCGGAGATCAGTGTGT 2388
 DB 31450 ATTGAGTTGATATTAAGATGCTACTTCTTTATGATGAATTAATGAATAATGTTGT 31509
 QY 2389 GTGCAAAAGGTGACTATGAGTTATCTGAGCGACGACGACGAG 2434
 DB 31510 TCTGAAGAGGGGAGTACACGTTCTTGTGGAACACGACTGAAG 31555

RESULT 6
 AF329731
 LOCUS 2793 bp mRNA linear PLN 28-JAN-2001
 DEFINITION Volvariella volvaria beta-glucosidase mRNA, complete cds.
 ACCESSION AF329731
 VERSION AF329731.1 GI:12584216
 KEYWORDS
 SOURCE Volvariella volvaria.
 ORGANISM Volvariella volvaria
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 Agaricales; Pleuroceae; Volvariella.
 REFERENCE 1 (bases 1 to 2793)
 AUTHORS Ding S., Ge W. and Buswell, J. A.
 TITLE Molecular cloning, characterization and expression of a
 beta-glucosidase gene from the edible straw mushroom, Volvariella
 volvaria
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2793)
 AUTHORS Ding S., Ge W. and Buswell, J. A.

TITLE Direct Submission
 JOURNAL Submitted (15-DEC-2000) Biology, The Chinese University of Hong
 Kong, Shatin, N.T., Hong Kong, China
 FEATURES
 source
 1. 2793
 /organism="Volvariella volvaria"
 /strain="V14"
 /db_xref="taxon:36659"
 41. 2629
 /EC_number="3.2.1.21"
 /product="beta-glucosidase"
 /protein_id="AAG59831.1"
 /db_xref="GI:12584217"
 /translation="MPPSDFAKANDEIVEQLTLDEALSTAGVFMHTAERLGP
 AVSDDPGNIGRNFHFMGTGPKCLPSSPALATMDPEVEVEGKLALPEAKLAAS
 LVLAFTSNIORNDLGRSFSESDPYGTLISASVNVKVGIGATIKHFGVDE
 DDROGDSIISERALEITLPLMLTOKTAAHMAITATNRYNGHVAEDPRLKQVL
 RNMWKIKGLMSDFQWYSVDHGTAKGLDLEMPGINKMTLVLNRTIARKLTPRDI
 KDRARVLELVKCAQAGPEILDGDEERTVELESDKLMLRIASEIVLKNVLP
 LKREGGAIRKIAVVGNAKAQVLSGGSAALASYSISPYDIKALAEHVEVFESE
 GARAYKTLPLEMDETETGERGNGTGTHTSDSMATLDOPFTAPRLVDTRTETS
 TSPKGITRMTMRILKGLKPREKDNFEFGIILAGRAKIAWDGOLVDNMTROBGE
 APFGSSQSEETGYILLKAKKKEIYETGVNPAADGDEDAIMNSQVRLGAEVA
 NMDLLSEAVKLASEADAVIAVGLNADETGNDRTALGLRDELEKAVKNSK
 TVVVTGAGSAILPMLDSVAAYVAMYLNAHNGDAIDVLEKQNSGLSLTFPKRL
 EDVPSHGHSSENGKVRVAEDLEFVGKHHNNHINIPLEPFEGSLTYTFESLSNQLSA
 PYDHTSSESLATLSTNTGTPTGSEVAOLQVYVPESTETHAPLOLRARKYKDL
 OGFETFEVLELDKXAVSYWNRQYTMVAVENVEYEIKVGNSSVSDLLVQRFVVKDG
 FERRGI"
 BASE COUNT 697 a 696 c 737 g 663 t
 ORIGIN
 Query Match 15.7%; Score 392.2; DB 8; Length 2793;
 Best local Similarity 50.3%; Pred. No. 2.6e-38;
 Matches 1258; Conservative 0; Mismatches 1168; Indels 75; Gaps 9;

QY 17 TTGAGGCCATCTTGAGAGAGCTCAACCTGGCGGAGAGGTGATCTGCTGCTATGCG 76
 DB 72 TTGATGAATTCGTCGAGCAATTAACCTGATGAGGCCATCACTGACCGCTGTGCG 131
 QY 77 ACTTCTGGCACAAGAGGCTCCCGAAGATGAGTCCCTGCTCCGCTTACGAGAG 136
 DB 132 GTTTCGGCATPACATGCGCATTTAGCGCTGCTGCTGCGAGCTGAAGGTAGCGAG 191
 QY 137 GCCCAACGCGCTAGAGAGGACCAAGTCTTCAATGCGCTGCGGCTGCTCCCTT 196
 DB 192 GACCAAAAGCGCATTCGAGCAACCACTTCTTCAATGAGGACCGCTGCAAAATGCTACT 251
 QY 197 GCGGACAGCTGCTGCTGCTCACTTCAACCAACTGCTCTCGAAGAGCGAG---TAAGA 253
 DB 252 CTTCAACCGCTCTTGGCGCAACTTGGAGACCTGMACTGCTCGAGAGGCTG;GGTGGAGCG 311
 QY 254 TGAATGGCAAAAGAGGCGCATGCAAGAGGCGCATGATGCCGCGCGGACATCAACA 313
 DB 312 TCTTAGCCCTTGAAGCTGAGCTGCGAGCGGCTTCTTAGTCTGCTCTACTCTCAACA 371
 QY 314 TGCACGCTCCCTCTCGGTGACGCTGCTGCAATGCTGATGAGATTCGTTCTGCTG 373
 DB 372 TCAAAAGAAACCCCTGGGAGGCGGCTAGTTTGCAGAGTTTCTCTGAGATTCCTTACT 431
 QY 374 CGGCGTTGGAGCTGCGGCTCTCATCCGCGCATTCAGAGCACTGAGTGCAGGCTACGA 433
 DB 432 CGGGTATCATTTTGGGCGCTCTACGTAAGTGTGTGAGAGAGGTGAATTTGGTGTACTA 491
 QY 434 TCAAGCACTTTTGTGCAATGATCAGAGAGAGAGCGGCGATGATGGGCGCATCACTCA 493
 DB 492 TCAACATTTTGTGGAAGCGACAGAGAGAGCGATGTCAGAGATACGATATATATATA 551
 QY 494 CGGAGCGGCTCTCGTGAATCTACGACCTCCGTTCCAGATTTGCTGTGCGAGACTGCC 553
 DB 552 GCGACGCTGCTCTCGCGGAATCTATCTACTCCCTTTATGTTACTCAAAAGTATGCTG 611

QY 554 AGCCGGGTGCTTATGACGCGCTACATGTCATGCGCTGTCTGTGACGAGAAC 613
 Db 612 CCCCATGGGCAATCATGACCGCTTCAACGAGGAGGAGCGCTTCATGACGAGAAC 671
 QY 614 CTAAATATCTTGATGGATGCTTGAAGAGATGGGTTGGATGGCTATCATGAGCG 673
 Db 672 CTTCTGCTCTACAGAGTCTTACGTAATGAGTGAATACAGGAGTGTATCATGAGTG 731
 QY 674 ACTGTACGGACATACAGTACACAGAGCGCTTGTGACGCGCTTCGACTCGAGATGC 733
 Db 732 ACTGGTTCGATGTAGAGCGCTGTATGATCAAGCGCGCTGATGAGACTCGAATATGC 791
 QY 734 CCGGACTCCAGCTTCCGAGGA---GAAACTCAAGTTCAAGTCTCCACAGGGAAGC 790
 Db 792 CTGGATCAATTAATGGGCTACCTTGTATGATCAACCGTATTAATCAAGCTCGAAC 851
 QY 791 CTTTATCCAGTATTAATGACAGAGGCTAAGGAGTCTTCACTTGTCAAGAGTGTG 850
 Db 852 TCACCCCGAGAGATCAAGAGACGAGAGAGTGTATGAGTGTAGTGAAGAGTGTG 911
 QY 851 CT-----GCTCCGAGTGAAGGAGAACGCGCCGAGACAGTGTCAACACACCCCG 904
 Db 912 CTCAAGAGCGCCAGAGATTCGATGGGATGGGAGAGGACGTACAGTCAAGCTTGAT 971
 QY 905 AAAGCGAGCTCTCTCGGGAAGTGGCAACGAGGCGATGCTGCTGAAGAGAGCA 964
 Db 972 CCGAGAGCTGCTACGCGGAGATCGCATCTGAGTTCATGCTCTACTCAAAATGATA 1031
 QY 965 ACACGTTCTGCGCT-----TGACCAAGAGAGAGAGAGAGCTGATTTGTGCGCCCA 1015
 Db 1032 AGCTTCTCCTCTTAACCTGAAGAGAGCGGAGATCAGAAATGATCGCTCTGTTGGTGA 1091
 QY 1016 AGCCCAAGAGCGGCATACACAGCGGAGAGGCTGCGGCACTGACCTACTACAGCAG 1075
 Db 1092 ATGCAAGAGCGAGTCTCTGTCGCGAGAGGCTGCGGCGCTGAGCGCTCATACTTCA 1151
 QY 1076 TCACCTCCCTTGAAGGCTCAGCAAGAGCTGAGAGCGCGCCATCATGACAC----- 1128
 Db 1152 TCTCACCTTGAAGGCTTGAAGGCGCGCTGAGAGCTGAGAGCTGAGAGCTCATCTCA 1211
 QY 1129 -----GTGGCGCTTACACACCGCTTCTCCATCTTGAAGGAGAGTGTGCTCAGC 1180
 Db 1212 GTGAGGCTCTGAGCGCTTAAACACTACCCACCTCGAATGGATCTGGAACCGAGA 1271
 QY 1181 CCGAGGCGCTCGCGGAGCGCTGAGGAGCTTCAAGCGCGCCCTGGTA----- 1231
 Db 1272 CCGGCAAGCTGAGTGGATGAGTGGACGTGGCACAACCGAGTCCGACGACAGCATGAGCG 1331
 QY 1232 CCGCTAACCGCAGCAGCATTTGACAGCTTCTTCAACCAAGAGAGATGACACTGTGG 1291
 Db 1332 CTCTTACCAAGCATTCATCGCTCCCGCATGTGTTGACGAACAGAACTCTTCATCAGCA 1391
 QY 1292 ACTACTACACCCCAAGGCGGAGACAGTGTGACGCGCAGCATGAGAGG---CAGGTCA 1348
 Db 1392 CGTCTGATCCCAAGGGAATTAACAAAGAGTGGAGCATGAGTGAAGGCTTACTGAAAC 1451
 QY 1349 CCGCGAGAGAGCTGACCTACAGAGCTGCGCTGCTCTGCGGACGAGCAAGAGCGT 1408
 Db 1452 CAAAGAGAGAGAGACCAACTTGGATTTGCTTATGCTGTGCTGAGGCTGAAGCTCT 1511
 QY 1409 ACCTAAGCAGCAGCTGCTGCTGACACGACCAAGAGGCTCCCGCGATGCTTCT 1468
 Db 1512 GGGTTTACGGAACATGTAATGACAACTGAGACCGCTCAACGCTCTGTGTAAGCATTT 1571
 QY 1469 TCGGCTCCGCGACCGCGAGGAGAGGCGGATCAATCTCGTCAAGGAGCAACAGTACA 1528
 Db 1572 TTGGCAGTGTAGTCAAGAGAGAACTGGGCTTATCTCTCAAGGCTGGCAAGAGCATG 1631
 QY 1529 AGTTCAAGATCGAGTTGCGCTC-----CGCAACCACTACACCTCAAGGCGGACACCA 1582
 Db 1632 AATCTATGTCAGTACTGTAAATGTGCGGCTCGGCTGATGTGATGAAGATGAGGCA 1691
 QY 1583 TCGTCCCGGCGACGCGCTCCCTCCGCTGCGGCTGCAAGGATGATTTGACGACGACGCGG 1642

Db 1692 TTAATGATAGTAACCTGTGTTCCGCTTGTGTGTCAGAGGTTGCCAACGAGTATC 1751
 QY 1643 AATTCAGAAAGTCCGTCGCGCTCCGCAAGAGAGACAGCAAGGATCATCTCGCGGCGC 1702
 Db 1752 TCCATCTGAGGCTGTCAAACTGCTTCAAGAGCTGATCCCTCATTTGCGGTGGTGC 1811
 QY 1703 TTAAGCCGACCTGAGAGACCGAGGCGCGACCGCGAGCATGAAAGCTCCGCGAGTGC 1762
 Db 1812 TCAACGAGATTTGGAGACTGAGGATTAATGACAGAGGACATTTGGCTTCCGCGGCA 1871
 QY 1763 TGAAGCACTATGTCGAGAGTGGCGCGCGCAACCCAAACACCGTGTCTCATGACA 1822
 Db 1872 CGGATGAATTAATGTAAGAGTGGCGAGGTCAACTGGAAGACTGTTGATGATCCAG 1931
 QY 1823 CCGGACCGCCGAGAGAGTCCCTGGCTGAGGCGCAACCGCGCGCATTCAGAGCTGT 1882
 Db 1932 CTGGTCACTATTAATCTTTCATGAGTGCATCTGTCTGATGAGTGTTCACGATGAT 1991
 QY 1883 ACAGCGCAAGAGAGAGGAGCTCAATGTCGAGCTGTCTTGTGCGACTACACCTT 1942
 Db 1992 ATCTTGTATGCTACTGAGAGATGATGATTCGAGAGCTGTGCGCAAGCAAGATCAT 2051
 QY 1943 CCGGCAAGCTGTCTCTGAGTCTCCCAAGCGCTCGAGAGCAACCCCGGTTCTCACT 2002
 Db 2052 CTGTGAATTTGCTGTGATCTCCCTAAGAGTGGAGAGCTGAGCTTCCATGAGACT 2111
 QY 2003 TCCGCAACGAGCGCGCGAGCTGATGAGAGAGTCTGATGAGTGTGAGTACT 2062
 Db 2112 TCGTTCGAGAGAGAGAGAGATGATGATGCGAAGACTTGTGTTGATCAAGACT 2171
 QY 2063 ACAGATTTGCGAGAGAGAGCTCAATTTCCCTTGTGCGACAGGCTGTCTACACACT 2122
 Db 2172 ACCACACCGAGAGAGAGAGCTTCTTCCCTTGTGCTGTGCTCTACACACT 2231
 QY 2123 TTGCTTTTCAATCTCTCCGT-----CTCACAGAGAGCGCA 2161
 Db 2232 TCTCCTCTCAATCTTCACTCTGCGACACAGTATGATCAAGCAACTCATCTTCT 2291
 QY 2162 AGCTGAGGCTGCTCTCTGCGTGAAGAACACCGGCTCGTGCAGCGAGAGTGGCC 2221
 Db 2292 CTCTACCGGACACCTTCTTCACTACCAACCGGCTCCGTTACCGGCTCGCAATCCCG 2351
 QY 2222 AGCTTACGTAAGCCCTTCCAGCGGCAAGATTAACCGCGCGTCAAGAGTCAAG 2281
 Db 2352 AACTTACGTAAGTATCCGAACTCTGGAATGACACAGCACTACAGCTGAGG 2411
 QY 2282 GCTTCGAAAGTGAAGTGAAGTGAAGCGCGGAGAGAGGAGTGAATGAGAGAGG 2341
 Db 2412 CTTTCAAGAGTGAAGAGACTTGAAGCTGCGAGAGACTGCGAGTGAAGTGAAGTGG 2471
 QY 2342 AGAATAGCTGCTGCTGATTTTGAAGAGAGCGGATCACTGTGTGAAAGAGGTG 2401
 Db 2472 ACAAATATGCTGTGTAATGGAATGACAGAGTATCAGAGCTGTGGCTGTGGAAGCGG 2531
 QY 2402 ACTATAGAGTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2442
 Db 2532 AGTATGAGATCAAGGTTGAAATGACAGCTGTCAAGAGAT 2572

RESULT 7

AF005277 3578 bp DNA linear, BCT 02-MAR-1998

LOCUS AF005277 3578 bp DNA linear, BCT 02-MAR-1998

DEFINITION AF005277 3578 bp DNA linear, BCT 02-MAR-1998

ACCESSION AF005277 3578 bp DNA linear, BCT 02-MAR-1998

VERSION AF005277.1 GI:2921739

KEYWORDS

SOURCE

ORGANISM

Cellulomonas blazoea

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Micrococciaceae; Cellulomonadaceae; Cellulomonas.

REFERENCE 1 (bases 1 to 3578)

Wong, W.K., Ali, A., Chan, W.K., Ho, Y., and Lee, N.T.

Db 1884 AGCTTACCTCTGGGACAGCCGACCCGACCGGAGCGGGCGCCGCCGAGCGCTC 1943
 QY 1594 CACGCG-----TCCCTCCGCTGCGCGGCTGCAGATCATATACACAGCGCCGAA 1644
 Db 1944 GCGGGGCTGGGCTGTGTGGGCTGTGCGCCGACCCCGGAGGCGGAGGACGACGTC 2003
 QY 1645 ATGAAAAATGCTGCTGCTGCTGCGCAAGACACGACAGGTATCATCTGCGGGCTT 1704
 Db 2004 ATCACCCTGGCGCGCGCGCGCGCGCGAGGCGGAGTGTGCGCTGTGCGGGGCTG 2063
 QY 1705 AACGCGAGCTGGAGGAGGCGCGCGCGCGCGAGATGATAGTCTCCCGCGCTGTG 1764
 Db 2064 ACCGAGAGAGGAGGAGGAGGAGTGTGCGCAAGTCCAGATGCGCTGCGCGCGCG 2123
 QY 1765 GACACGATATTTCCGACGTGCGCGCGCGCGGAAACCAACCGCTGCTGCTGAGAG 1824
 Db 2124 GACGCGCTGTGCG 2183
 QY 1825 GGCACCCCGGAGAGATGCCCTGCTGCAAGCCGCGCGCGCGCGCGCGCGCTGTAC 1884
 Db 2184 GCGACCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2243
 QY 1885 GCGGCGAAGAGAGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
 Db 2244 CCGGCGAAGAGGCG 2303
 QY 1945 GCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
 Db 2304 GCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2363
 QY 2005 GCGACCGAGCTG 2064
 Db 2364 CCGCTGTGCTG 2423
 QY 2065 GACTTGGCGAGAGAGAGAGTTCCTTCCCTTGGCGAGCGCTGCTGCTGCTGCTG 2124
 Db 2424 GCGGACCGGCG 2483
 QY 2125 GCTTTTCCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2184
 Db 2484 GACTAGCGCGAGCG 2543
 QY 2185 AAGACACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2244
 Db 2544 ACCAACAACGCTG 2603
 QY 2245 GCGGCGAATTAACG 2304
 Db 2604 TCCGAGGAGCTG 2663
 QY 2305 CCGCGGAGAGC 2316
 Db 2664 GCGCGGCTGACG 2675

RESULT 8
 ATUCBG1 3710 bp DNA linear BCT 26-APR-1993
 LOCUS A.tumefaciens beta-D-glucosidase (cbg-1) gene, complete cds.
 DEFINITION M59852
 ACCESSION M59852.1 GI:142221
 VERSION beta-D-glucosidase.
 KEYWORDS A.tumefaciens DNA.
 SOURCE Agrobacterium tumefaciens
 ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium.
 REFERENCE 1 (bases 1 to 3710)
 AUTHORS Castle, L.A., Smith, R.D. and Morris, R.O.
 TITLE Cloning and sequencing of an Agrobacterium tumefaciens beta-glucosidase gene involved in modifying a vir-inducing plant signal molecule
 JOURNAL J. Bacteriol. 174 (5), 1478-1486 (1992)

MEDLINE 92165721
 PUBMED 1537792
 FEATURES
 source location/Qualifiers
 1..3710
 /organism="Agrobacterium tumefaciens"
 /db_xref="taxon:358"
 /cell_line="B3/73"
 /tissue_lib="genomic HindIII cosmid"
 169..2637
 /gene="cbg1"
 169..173
 /gene="cbg1"
 181..2637
 /gene="cbg1"
 /EC_number="3.2.1.21"
 /trans_start=1
 /trans_table=1
 /product="beta-D-glucosidase"
 /protein_id="AA22082.1"
 /db_xref="GI:142222"
 /translation="MIDDIIDKMTLEQVSLSGADEFTVAERLGVPRKITYDGPNGARGSLVGVKSPAPVALALGATWDPDLERAGVALGQKSGASVLAFTPIHRSGLNRFECVSEDPALTAACAVATINOSQVATIKHPANSELEROTMSDVDETLREITLPPPEAVKAGKAVMSYKLNKTYTSENPMILKTVLREKPPGCVVMSDFESHSTAETINAGLDLEMPWRDGEKLVAAVRRGKVAETVRASARILLLEKRGAREKAPDLAEHALDLEPDRALIRQAGARAVILKNDGVLPLAKSFDIIVAGPNAASARVMGGSARIAHAYVSPLEGRALSNANSLRAHVCNNRRLIDVSEGMTVEYKRGFESRPVHVEYKGFEMFDLPISGLDLADFSARMTAFVPEQTEHIEGKASIDGINITALREYKPLIGDNLAEAVETARKSDIVLLVREGEMDEGLDIDPLRPGROBELLEAVETNPVNVVLTQGGPIEMFWLKVRAVLQMWYPGDGLNALDYLFGVEPAGRLPOTFKALDNSITDDPSITPGODHVRVAGIIVGYNHHTREPLEFPGELGVTRETWGAPOLSGEMGADLTIVDVNTNGDRAGSDVOLLVHSPNARVEPFEKELAPFKLAPGATGAVILKINPDLAVFDVEAGHPRADACKYELIVASAIIDRASVSHLPVDHMEP"
 213..223
 misc_feature
 /gene="cbg1"
 /note="active site"
 BASE COUNT 718 a 1084 c 1168 g 740 t
 ORIGIN
 Query Match 13.7%; Score 343.4; DB 1; Length 3710;
 Best Local Similarity 49.1%; Pred. No. 1,7e-32;
 Matches 1195; Conservative 0; Mismatches 1171; Indels 69; Gaps 8;

Db 17 TTGAGGCATCTTGAAGAAGCTACCCCTGCGGAGAGTGCATCTGCTGCTGCTGCTG 76
 Db 185 TCGAGATATCTGATTAAGTACACTGAGAGACAGGTCTGCTGCTGCTGCTGCTG 244
 QY 77 ACTTCTGACACAAGAGCTCTCCCAAGCATGAGAGTCCCTCTCTCCGCTTTACAGATG 136
 Db 245 ATTCTGTGAGACGCGTGCATGACAGCGGCGTCCGCGCCCAAGATCAAGTTACGACG 304
 QY 137 GCCCAACGCGCTAAGAGGACCAAGT---TCTCAATGCGCTGCTGCGGCTGCTCC 193
 Db 305 GCCCAATGCGGCAACGCGCGCGCGCGCGCGCTGCTGCGCGCTGCTGCTGCTG 364
 QY 194 CTTGGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 253
 Db 365 CCGTGGCATATGCGCTTGTGACCGGACGCGGAGCTATGACGCGCGCGCGCTGG 424
 QY 254 TGATGGCAAGAGGAGCATGCTAAGAGTGGCATGTGATCTGCGCGCGCATTAACA 313
 Db 425 CGTGGGAGGAGCAAGCAAGAGGCGCGCGCGCTGCTGCGCGCGCATTAACA 484
 QY 314 TGCACGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 373
 Db 485 TTCAACGCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
 QY 374 CCGGCTTGGAGAGTGGCGCTCTCATCCGCGCATTCAGAGACATGAGAGTCCAGCTCA 433
 Db 545 CCGCGCGCTGCGCGCTGCGCTATATCAATGCGCTGACAGACGAGGCTGTGCGCGCA 604

QY 434 TCACGACCTTTTGTGCAATGATCAGAGACAGGCGCATGATGTCAGACATCGCA 493
 DB 605 TCACGACCTTGTGCGCCAGAGTCGAGATCGAGCGCACAACCATCTCTCCATGTGC 664
 QY 494 CGGAGCGGCTCTCCGTGAATCTACGCACTCCGTTCCAGATTGTGTGAGACTCC 553
 DB 665 ATGACCGGACGCTGGCGCAAAATCTATCTGCGCCCTTTGAGAGGCGGTGAAGAGCGC 724
 QY 554 AGCCGGGTGCTTCACTGACGCGCTACATGCGATCAATGGGTGTGTGACGCAAGAAC 613
 DB 725 GCGTAAAGCGCGCTCATGTCTCTCTACAAACAGCTCAACGACCTATACGAGAAATTC 784
 QY 614 CTAATATCTTGTAGTGGATGCTTGAAGAAATGGGTTGGATGGGCTATCATGACG 673
 DB 785 CCTGGCTGTGACGAAATGCTTGGCGGAGAAATGGGCTTGGACGCGCTGGTCAATGTCC 844
 QY 674 ACTGTACGCGCATACAGTACAGTACAGAACCGTTGTGAGCGCTGACCTCGATGTC 733
 DB 845 ACTGGTGGCTCGCAGTCGACGCGTGAACCATCAATGCGGCGCTCATCTGGAGATGC 904
 QY 734 CCGGACCTCCACGCTTCCGAGGAGAAACATCAAGTTCAAGTCTCCAAAGGAAAGCCCT 793
 DB 905 CCGGACCTTGGCGGAGTGGCGGAAAGCTGTGCGCGCATGTCGGAGAGCAAGTAA 964
 QY 794 TTATCCAGCTCATGTACGAGAGGCTAGGAAATCTTCAATCTGCTCAAGAAATGTGCTG 853
 DB 965 AGCGCGAGACCTGTGGCGCTTCCGACAGCGATATCTCTCTGTGCTGACAGCGCTGGCG 1024
 QY 854 CCGCGAGTGAAGAGAGAGAGCGCCCGGAGAGACTGTCAACAAACCCCGAAAGCGGAG 913
 DB 1025 CTTTGAAGAGCGCTGTATCTGCGGCAACAGCGCTGTATCTGCGGAGAGTGTGCGC 1084
 QY 914 CTTCTCTCCGGAAGGTGGCAAGAGGCGATGCTGTGTAAGAGAGAAACAACGTTTC 973
 DB 1085 TCATCCGCAATCTCGTGTGGAGGTGCTGTACTCTTAAGATATACGATGCTGCGCC 1144
 QY 974 TCCTCTTGAAGAGAGAGAGAGAGAGCTGATTTGTGCGCCCAACGCAAGAGCGCAT 1033
 DB 1145 TTGCGCAAGTGTCTTGTGACAGATCGCGCTCATCGGCGCCCAATGCGGCTTCCGACG 1204
 QY 1034 ACCAGCGCGAGAGCTTGTGCGCACTCAGAGGCTTACTAGCAATCACTCCCTTTGAGCGC 1093
 DB 1205 TCATGGCGGAGAGAGCGCGGATGCGCATTTATACGTTGAGCCACTTGTGAGGCA 1264
 QY 1094 TCAGCAAGAGCTGAGAGAGCGCGCATGTAACGCTGCGCGCTACACACAGCTTCTCTC 1153
 DB 1265 TTGCGCGCGCGCT-----GTCCAGCGCAACAGCTTCCGCGCATG 1303
 QY 1154 CCATTTCTAGGAGAGAGAGCTGCTACAGCGCGCGCTCCGCGCATGCGCTGAGAGGCT 1213
 DB 1304 CGGTGCGGTGTAAACAACAGCGCTCATGCAAGCTTTCAGCGGAGATGAGCGGTGCAAT 1363
 QY 1214 TCACGAGAGCGCGCTGTACCCCTAACCGCGACATGACATGAGAGCTTTTTCACCAAGA 1273
 DB 1364 ACTTCAAGGAGCGCGCTTGAAGAGCGCTGCTCATGTGAGACCGCTGAAAGAGCGC 1423
 QY 1274 CGGACATGCACTGTGAGACTACTACACACCCAGCGCGACAGACAGCTGTACGCGACA 1333
 DB 1424 AATTTCTTGTGTTGATCTTCCGTCCGCGAGACTGTATCTGCCCATTTTTCGGCGCA 1483
 QY 1334 TGGAGGCACTACACCGCGAGAGAGACTGCACTAGAGAGCTGCGCTGCTGTGCTGCG 1393
 DB 1484 TGAAGGCACTTGTGCGCGAGAGAAACCGGTGAACACATCTTGGCAATGACATGTG 1543
 QY 1394 GCAAGGCAAGAGCTACAGTACAGACAGCTGCTGTGACACAGCCCAAGCAAGGTC 1453
 DB 1544 GCTTGTGCTGCGCTGTGAGAGCGGCACTGTGTGATGATGATGACGCTTGGAGCA 1603
 QY 1454 CCGGAGATGCTTGTGCGCTCCGCAACCGCGAGAGAGAGCGCGCGCAATCATCTGCA 1513
 DB 1604 AAGGTGAGACTTTTGTGAACCGCGAAGAGAGAGAGGCTGCGCGGTAAGCTTGGGG 1663
 QY 1514 AAGGCAACAGTAAAGTTCAAGATCAGATGCGGCTCCGACCCACTACACCTCAAG 1573

DB 1664 CCGAGCGCCCTACCGGGTGTGTGCAATATG-----AGGCGCCGAAG 1708
 QY 1574 GCGACACATCTGTCCCGGACAGGCTCCCTCCGCGCTGCGGCGCTGACAGTCAATGAG 1633
 DB 1709 CCAAGCTGAGCGGCTACATATATGTGCTGCTGCTGCTGCTGCAAAAGC-35CT--CG 1765
 QY 1634 ACCAGCGCAATGCAAAATGCGCGCTCCGCAAGAGACAGACAGCTCATCATCT 1693
 DB 1766 GCGATGCGGGATTTGCGAGAGCGGTGAAMACCGCCCGCAAGTCCGATATGCTGCTC 1825
 QY 1694 GCGCGGCTTAAACCGGACTGAGAGACGAGGCGCGCCAGCGGAGATGAACTCC 1753
 DB 1826 TCGTGGCGCTGAGGCGAGATGGAGACCAAGGCTGTGATCTGCGGACATGCGCTCC 1885
 QY 1754 CCGGCGTGTGAGACAGCTATTTGCCAGAGTGGCGCGCGCGCAACCCAAACCGCTGCG 1813
 DB 1886 CCGTCCGAGAGAGATGATGAGCGGTGCGCCGAACCAATCCCAACGCTGTGCTG 1945
 QY 1814 TCATGACAGAGCGGACCCCGAGAGATGCTGCTGACAGCGCACCGCGCTCATCC 1873
 DB 1946 TACTGCAAGAGGGTGTGCTCATGAGATGCTGCTGCGCAAGTGGCGGTGCTGCTG 2005
 QY 1874 AGGCTGTGAGCGGCGCAAGAGAGAGCGGCACTTCATGCGGAGCTGTGCTGCGACT 1933
 DB 2006 AGATGTGATATCCCGGCAAGAACTTGGCAATGCGCTGTGGAGAGCTTCTTGTGATG 2065
 QY 1934 ACAACCCCTGGGCAAGCTTCCCTGAGTTCCCAAGCGCTGACAGCAACCCCGCT 1993
 DB 2066 TCGAGCTGCGCGCGCTTCCACAGACTTCCCGAAGGCGCTCAGAGATATTCGCGCA 2125
 QY 1994 TTCTCAACTTCCGAC-----CGAGCGCGGCGCACAGCTGTACGCGGAGAGC 2041
 DB 2126 TTACGACGATCCGCTCATCTATCTGCTGAGAGAGCGGCTATGCGCGTACCGAGAGGA 2185
 QY 2042 TCTACGCTGCTGAGTACTACGATTTGCCAAGAGAGCGCATTTCCCTTTGGGC 2101
 DB 2186 TCTTGTGCTGCTATCGCATACGATACAAAGAGATCGAACCACTTCTTCTGCGCT 2245
 QY 2102 ACAGCTGTCTTACACTTTTGTGCTTCAATCTCTGCTGTCTACAAAGA----- 2156
 DB 2246 TCGGCTTGTGCTACACCGGCTTTTACCTGAGGCTGCGCGCAAGTATGGGAAAGGAATG 2305
 QY 2157 ---CGCAAGCTGAGCGTGTCCCTCTCGTGAAGAAACACCGGCTCCGCGCGCAC 2212
 DB 2306 GCGGAGTGTGCTTACGATGACGCTCATCTACCAATATAGCGACAGGCGGATCGG 2365
 QY 2213 AGTGGCGGAGCTTACGTAAGGCGCTTCCAAAGCGGCAAGATTAACCGCGCTCAAG 2272
 DB 2366 AGTGTGTGAGCTTATGTACTCTCCAA---TGCAGAGTTCAGAGGCTCCTTCAAG 2422
 QY 2273 AGCTCAAGGCGTGTGCAAGGCTGCAAGCTGACCGCGGAGAGAGAGGCTGTAATCG 2332
 DB 2423 AGTGTGCTGCTTGTGCAAGCTGCAAGCTGCGCGCGGCGGAGCGGATGAGC---CGGTG 2479
 QY 2333 AGGAGCAGAGAGATGCTGCTGCTGATTTTGTATGAGAGCGGAGTCAATGTGTGTC 2292
 DB 2480 TGAAGATGCTCTCGGAGACTTGTGCTTACTTGTGATGAGCGCGGCTGCTTCCGGCTG 2539
 QY 2393 AAAAGGTGATATGAGTATCTGTAGCGACAGC 2427
 DB 2540 ATGGGCGCAATGACGACTATCTGTGCGGCGACG 2574

RESULT 9
 ABI293760 2636 bp mRNA linear pln 09-AUG-2000
 LOCUS
 DEFINITION
 Agaricus bisporus mRNA for putative beta glucosidase (bgl1 gene),
 strain D49.
 ACCESSION
 AJ293760
 VERSION
 AJ293760.1 GI:9796011
 KEYWORDS
 beta glucosidase; bgl1 gene.
 SOURCE
 Agaricus bisporus.

ORGANISM Agaricus bisporus
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 Agaricales; Agaricaceae; Agaricus.

REFERENCE 1 (bases 1 to 2636)
 TITLE Molecular analysis of the cellulolytic genes in Agaricus bisporus
 AUTHORS Morales-Almora, P. and Thurston, C.F.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2636)
 AUTHORS Morales-Almora, P.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-2000) Morales-Almora P., Microbiology, King's
 College London, 150 Stamford Street, London, SE1 8WA, UNITED
 KINGDOM

FEATURES
 source
 1. .2636
 /organism="Agaricus bisporus"
 /strain="D649"
 /db_xref="taxon:5341"
 /clone="PMR1615"
 /country="United Kingdom"
 25. .2610
 /gene="bg1"
 25. .2610
 /gene="bg1"
 /EC_number="3.2.1.21"
 /function="putative hydrolysis of beta-D-glucose"
 /codon_start=1
 /product="putative beta glucosidase"
 /protein_id="CAC03462.1"
 /db_xref="gi:9796012"
 /translation="MPPSDPANADIDKIVDALTTDEAILLSAGVGFHMTAIERLOIP
 AVKSDGPNIGRNHFMFGTAKLPSTAMATPDRDLEEGKLRLKREKRSAS
 VILAPNCINIDRLGGSPESFSEDFSLGMLGSAITNGKJOKGJATIKHEFNKKE
 DHRGYSIHRTSSGDLSPSVHSAKRNPLHGAIVATNPNRNGIHVCEENPELLOKI
 LRDEKMSDVAHNSDMRTYSVDYGLNAGDLEPRGLAKRSLSESVRSIQSKRYTRKK
 VKERARVLELVKCAOGAPEILDGDELTDSDDNLMKFEAESIVLKNDRKV
 LPLDRSLKRLVAIVGNANAAVLSSGGSAAKRSYVARPAGLVKALKENPNIEVTV
 SEGARAKTLPDLEFTESGRGWTSVHAHENDSLVALPTPIKTOYVETRMFI
 SSSPAGITKRTLRMKGYLKPREDCKEFGGLAARAKLEVDGKLVINDNTRORG
 VEEFGGSEETGVELKAGVKEIIVDRCNVGPPADGETETIMSGEVRIRLGAEEV
 EDDDELINSVSIKADADAVIIVYGLANGMETEGNRTTIALPRDELIVOKYVAMP
 RAYVTDAGSALTIPWDOVSLHTWYLSNATGDAIDVIEGNHPSKLSLTFEKR
 LEDVPAGHYHSEKRVYLAEGLYVGYKHQHNRIEPLAFHGLISYTSFENLSDLS
 OPTVSGDFDLKATVRLKINTGNITGSQVVOYLGLPKTSELIRPMOLGFKKMRDK
 PGESREVELYMDRLSVSYMDKEMWVENGAADVAVTSEBGVEGEGELLGRFNEVKEG
 GMRGL"

BASE COUNT 657 a 603 c 682 g 694 t
 ORIGIN

Query Match 12.1%; Score 302.2; DB 8; Length 2636;
 Best Local Similarity 49.8%; Pred. No. 1.6e-27;
 Matches 1100; Conservative 0; Mismatches 1048; Indels 59; Gaps 11;

QY 2 TGGCTGATATTGATGTGAGGCCATCTTGAAGAAGTCACCCCTGGCCGAGAGAGTGTCATC 61
 DB 41 TTGGCAACGCGTGCATTTGATTAATTTGTCAGCATTTGACACGATGAGAGCCATTGTTT 100

QY 62 TGGCGGCTGATGACATCTGCGACACAAAGGCTGCCCAAGCATGAGAGTCCCTCTC 121
 DB 101 TGAAGTCCGAGGTGATTTTGGCATCTCTGCAATTGAGAGGCTGCAAGATTCCCGCTG 160

QY 122 TCCGCTTACAGATGAGCCCAACGCGTAAAGAGGACCAAGTTCTTCAATGAGCGCTCCCTG 181
 DB 161 TGAAGTTAGGAGAGGCGCCCAATGATCGAGAGAAATTCATTCTTTATGAGGCGAGCGCAG 220

QY 182 CGGCGCTGCTCCCTGGGCGACGTGCTCGCTTCCACATTCACCAAACTCTGCTGCAAG 241
 DB 221 CCAAGTGTTCGCTCATCTCCTCAATGAGCGCAACCTTTCGATCGAGACTTACTTGAAG 280

QY 242 AGGCAAGT---AGATGATGGGCAAGAGGCGCATGCTAAGAGAGTGGCGCATGATCTCTCG 298
 DB 281 AAGTTGGTCTCAAGCTTCTCGCTAGAGAGCAAGCTCAGATCCGCTCTGTTATCTCTCG 340

QY 299 GCCCGACTATCAACATGCAACGCTCCCTCTCGGTGAGACGTGGCTTCGATCGATTTGGTG 358
 DB 341 CTCCACTGTATGACTATCAACGATTAATCCCTCGGTGAGAGTTTGAAGATTTTCGG 400

QY 359 AGATATCCGTTCCCTGGGGGCTTGGAGCTGCGGCTCTCATCCCGGCGATTTAGAGACATG 418
 DB 401 AGAGCCCTTTCCTGCTGCTGATCTTTGGAGTCCCTACCAACGGAATTCAGAAAGGGG 460

QY 419 GAGTCGAGGCTACATCAAGACACTTTTGTGATGATAGAGAGAGAGAGAGAGAGAGAG 477
 DB 461 GTATTGGCGGCGACTATCAACGATTTGCTGGGAATATACAGAGAGATGATCCAGCTGGAT 520

QY 478 GTCCAGAGCATGCTCAACGAGAGCGGCTCTCCGTAATCTACGACATCCCTTCAGATT 537
 DB 521 ATGACAGTATTATTTACGAGAGACCTCTCGGAGATTTATCTCTCTCCCTTCATATGCTG 580

QY 538 GCTGTGCGAGACTCCCAAGCC--GGTGCCTTATGACGCGCTACAAATGATCATGATGCGG 595
 DB 581 GCTCAAGAAATATGCTCTTCATGAGGCGCATGACCGCGTACAAACCGTGTAAAGGATA 640

QY 596 TGTGTGACGAG 655
 DB 641 TTACAGCTTTCGAGAAATCCCGAACTTCTACAGAAATATTACGCGATGAGTGAATCAG 700

QY 656 ATGGCTTAATCATGACGAGCTGTGACGACATACAGATCAAGTCAAGAGAGAGAGAGAGAG 715
 DB 701 ATGTCATGATCATGACGAGCTGTGACGAGCTGTGACGAGCTGTGACGAGCTGTGACGAG 760

QY 716 GCTTCAGCTTCGAGATGCGGCGGAGCTTCACAGCTTCG---AGAGAAACATCAAGTTCA 772
 DB 761 GTCTTACCTTGAAG 820

QY 773 AGCTTCCCAAG 832
 DB 821 CAATTCAATCTAGAGAAATGATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880

QY 833 AGTTCTCAAG 886
 DB 881 AGCTTCAAG 940

QY 887 CTGTCAACAAACAG 946
 DB 941 GCACCTTGTATGAG 1000

QY 947 TGGCTGGAAG 1000
 DB 1001 TTCTGCTCAAG 1060

QY 1001 TGAATGTGCGCCCAACAG 1060
 DB 1061 CAATCGTGGGTGATGATGCAAG 1120

QY 1061 GGGCTCTACAG 1120
 DB 1121 AGCGGTCTTACCTGTCGCGGCGGCGGCGGCGGCTGTGAAAGAGATCAAC 1180

QY 1121 CGTACACGCGGCGGCGCTACAC-----ACGTTCTCTC 1154
 DB 1181 CAACATCATGAGTTACTTACAG 1240

QY 1155 CATTTAG 1214
 DB 1241 ATTGATTTGTTCACTGAGTC--CGGCCAAAGCGGTGAGAGAGAGAGAGAGAGAGAGAGAG 1299

QY 1215 CAACGAGCC---CCCTGTACCCCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
 DB 1300 GAAATAG 1359

QY 1271 AG 1330
 DB 1360 ACAAGATGTCTATGACAGAGTTCTCTCGCGGATATACCAAGAGAGAGAGAGAGAGAG 1419

QY 1331 ACATGAG 1389

QY 1147 GTTCTCCCATCTAGGCGAGCACTGCTCAGCCGCGGCGCTCCGGCATCGCGTGG 1206
 DB 80015 TATC-----GGTTTCAGCCGCTAATCGAAACCCACCACTTCGAGTTTTCACAGG 79963
 QY 1207 AGGCTTTTCAACGAGCCCGCTGATCCCTAACCGGACACATTGACGAGCTCTTCTTC 1266
 DB 79962 AAGGAGCTAGCTGCGCAACCGG-----GAAAGTCTGCGAGGAAACGCTGAGC 79914
 QY 1267 ACCAAGAGGACATGACACTGTGTGACTACTACCAACCCCAAGGCGGACAGACGTGTAC 1326
 DB 79913 CTGTGTATGATGTGCCCCCGGCGAGAGGCTTTGTGATCCCTTCGCG---TTTTCG 79857
 QY 1327 GCCCAATGAGAGGCGAGTACACCGCGAGAGAGTACGACATGACGCTGCGGCTGCTG 1386
 DB 79856 GCTAAGTGGCCACCACTTTCACAGCTCGAAGCGCGCTGTATCGCTGTGAGATTAAC 79797
 QY 1387 GTCGCGGCAAGGCGGAGGCTAGCTAGACAGACAGCTGCTGTCGCAACAGCCACCAAG 1446
 DB 79796 TCGGCGCGGCTTGGCGCGGTCTATGTGACGCGAGACTTGTGTGAGCGCTTGGGCTTC 79737
 QY 1447 CAGTCCCGCGGAGTCTTCTGCTCGCCACCGCGGAGAGACGGCCCGATCAT 1506
 DB 79736 TGGATAGCTGCTACCACTTCTTGAAGAAAGGCTGCGAAGAGCTGTGGTGAATACG 79677
 QY 1507 CTGCTCAAGGCGCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1566
 DB 79676 CTGCAAGCGCGCTGATACATGAGTGTGCTGCGAGTACGCGGACGACAGCTAC 79617
 QY 1567 CTCAAGGCGCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1626
 DB 79616 CTCAATATGCGCGCAATTCGGGTAGGAATAGC-----AGA 79581
 QY 1627 ATTGACAGACAGCGCGAATGAAAAGTCCGCTCCGCGCAAGGACGACAGCTC 1686
 DB 79580 TTTTCGCGCGGAGGAGATGCGGAGGACAGCAGCTGCGGCGCAAGGCTGATCAATGCG 79521
 QY 1687 ATGATGCGCGGCGCTTAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1746
 DB 79520 GTAGTCTTCTGCGGAGAGCGGCGGAGTATCCAGAGGCTCTGATCTTCCGCGTAT 79461
 QY 1747 AAGTCCCGCGGCTGTGAGACAGCTCATGCGAGCTGCGCGCGCGGCGGAAACCAAC 1806
 DB 79460 GCACATCCCGGCTGTGAGAAACAGCTGTGAGCGGTAATGCGGCAATCCCAACAGC 79401
 QY 1807 GTCTGCTGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1866
 DB 79400 ATCTGCTGATGACAAACCGGTGAGCGGTAAGTCCGCTGCTGCTGCTGCTGCTGCA 79341
 QY 1867 GTATCCAGGCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1926
 DB 79340 GTACTGCAATGCTGTATCCCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 79281
 QY 1927 GCGGACTACAAACCCCTGCGGAGAGCTGTCTCTCACTTCCCAAGGCGCTGACAGAAC 1986
 DB 79280 GCGAAGGCGGAGCGCTGCGGCGGAGTGTGACAGACTTCCCGCTGCGGCTGAGAAC 79221
 QY 1987 CCGCGGTTTCAACTTCCGAGCG-----AGCGGCGGCGGAGCGCTGTACGCG 2034
 DB 79220 CCGACATGACTGAGATGAGCGCGCTATATCAGGTAAGGAGCGGCGCTGTGATGAC 79161
 QY 2035 GAGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2094
 DB 79160 GAGGAGTGTGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 79101
 QY 2095 TTTGGCGAGGCGCTGTCTCTACACACTTTTCCCTTTTCCAACTCTCTCG-----TGTCT 2148
 DB 79100 TTTGCTGATGCTGTGCTGCTTATTAAGCTTTCGATGCTGAGGAGGAGGAGGAGGAGG 79041
 QY 2149 CACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2208
 DB 79040 GACGCTGCGCGGCGGTGACAGTGAAGTGAATGCAACATGAGGAGGAGGAGGAGGAG 78981

QY 2209 GCACAGTGGCGCCAGCTTACGTCAGAGCCCTTCAAGCGGCGCAAGATTAACCGCCGCTC 2268
 DB 78980 TCGGCACTGTGCAATATCTATCTCGGCGATGTTGAAGATC---TATACCGAGCGGCTT 78924
 QY 2269 AAGGAGCTCAAGGCGCTTCCGCAAGGTCGACACTGCAAGCCCGGAGAGA 2314
 DB 78923 AAAGAGCTGAAGCGCTTCTCGAAATTTGCCCTGGAACCGGAGAGA 78878
 RESULT 11
 AX489445 2601 bp DNA Linear PAT 16-AUG-2002
 LOCUS AX489445
 DEFINITION Sequence 6745 from Patent WO02053728.
 ACCESSION AX489445
 VERSION AX489445.1 GI:22323457
 KEYWORDS
 SOURCE
 ORGANISM
 Candida albicans.
 Candida albicans.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic Saccharomycetales; Candida.
 REFERENCE
 1. Roemer, T., Jiang, B., Boone, C., Bussey, H. and Olsen, K.L.
 Gene disruption methodologies for drug target discovery
 Patent: WO 02053728-A 6/745 11-JUL-2002;
 Elitira Pharmaceuticals, Inc. (US)
 FEATURES
 source 1. 2601
 location/Qualifiers
 BASE COUNT 917 a 334 c 484 g 866 t
 ORIGIN
 Query Match 8.2%; Score 205; DB 6; Length 2601;
 Best Local Similarity 46.7%; Pred. No. 7.6e-16;
 Matches 1031; Conservative 0; Mismatches 1097; Indels 78; Gaps 9;
 QY 11 TTGATGTTGAGGCACTTGAAGAGCTCACCCGCGAGAGAGTGTGCTGCTG 70
 DB 14 TTGATTTGATTAATATTAATATCAATTAAGAAAATAATGATTAAGTGGTG 73
 QY 71 GTATGACTTGTGGACACAAAGCTCTCCCAAGATGAGTCCCTCTCCGCTTGA 130
 DB 74 GAATTAATTTTGGCATATCATATCTTCGATTAACATTCGCAAGTGAATTA 133
 QY 131 CAGATGGCCCAAGCGGTAAGAGGAGCAAGTCTTCAATGCGCTGCGGCTGCT 190
 DB 134 CTGATGTCCCAATGATTAAGGAGCAAGATTTTCAATGATGCTCTGCGCTGTT 193
 QY 191 TCCCTTGGGCGACGTGCTGCTCCATTCACCAAACTGCTGCAAGAGCGAGTA 250
 DB 194 TCCCTGTGGAATGATTAAGTGTGCTGCTGATTAATTAATTAATTAATTA 253
 QY 251 AGATTAATGGAAGAGGCGCATGCTAAGAGTCCGATGATCCGCGGCGGACTATCA 310
 DB 254 AATTAATGAACATTAAGGCTAAATTAATTAATTAATTAATTAATTAATTAATTA 313
 QY 311 ACATGCAAGCTCCCTCTCTGCTGAGCTGAGCTTTCAGTGTGATGAGAGTCCGCTTC 370
 DB 314 ATATTCAGAGGCGCATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 373
 QY 371 TGGCGGCTTGGAGCTGCGGCTCTCATCCGCGGCAATTCAGAGC---ACTGAGGTGAG 427
 DB 374 TAACGTGCAAAATTTGCGAGTGTATTAATTAAGGCTATCAATTAATTAATTA 433
 QY 428 CTAGAGTACAGCACTTTTGTGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 487
 DB 434 CCAAGGTTAAACATTAATGTTGTAATGATTTAAGATGAAGAGAGTGTGATGATTTCT 493
 QY 488 TCGTTCAGGAGCGGCGCTCTCCGTAATCTACGACACTCCGCTTCCAGATTCGTGCGAG 547
 DB 494 TGGTTACACCAAGAGGATTAAGAGAAATTAATTAAGACATTAATTAATTAATTA 553
 QY 548 ACTGCCAGCGGCGGTGCTTATGACGCGCTGATGAGGAGTGAATGAGGCTGTCTGACGCG 607

D	554	AAAGTAATCCAAATTTGGTTTGGATGACAGGATATATAAAGTTAAAGTAAAGGTAACATGTTTCTC	613
O	608	AGAACCCTAAATATCTTGATGATGGATGCGTTGGAAAGAAATGCGGTGGCATGGCTTAATCA	667
D	614	AAAGTAAATTTTCTTCGCAAAATATATATACGTATGAATGAATTTGGCAAGGACAAATTA	673
O	668	TGAGGACTGGTACGGCCACATACAGTACCAACAGAAAGCGGTTGGCGAGCGCTCGACCTCG	727
D	674	TATCTGATTTGGTATGGGACATATACAGATGTAATAAAGCCATGTAAATGGTTGGANTTTGG	733
O	728	AGATGCCCGGACCTTCACAGCTTCGAGAGAAACACTCAAGTTCAACGTCCTCCAACGGAA	787
D	734	AAATCCCGAGATCTCCAAATTTTCGTATATAAACAATTTGTGACTAGTATGATTAATCA	793
O	788	AGCCCTT---TATCCAGTCAATTCACCAAGGGCTGAGGAAGTTCTTCAGTTCTGCTCAAG-	843
D	794	AAGATTCATATCAATCAACATTTGGATGATCTGTAAAGATGTTTTTAAATTAATTAATTA	853
O	844	-----AAGTGTGCTCGCCTCGGAGTACGAGAGACGGCCCCGAGACGACTGTCAACACA	898
D	854	TTGCCAAACAGTCAAGTGTGTGGTCACTGAAGATGTTAAAGAAATCATCGGAATTAACA	913
O	899	CCCCGAAACGGGACGCTCTCTCCGGAAAGTTTGGCAACGAGGCACTGCTGCTGGAAGA	958
D	914	CTCAAGAAACAAGAGATTAATCTACGCAAACTACCTCAAGATTCGAATGTTTTATTAAGAA	973
O	959	ACGAGAACAAAGTCTCGCCTTGTGACAAGAG-----AAGAAGACGC	1000
D	974	ATGATTAATTAATTTTATTCATTTGAAAGGATACGTCCAGCTTCAGCTCCAACTATTTG	1033
O	1001	TGATTTGTGCCCCCAACGCCAAGCCACATACACGGCGGAGGCTCTGCCGACATCA	1066
D	1034	CAATTAATTTGGTCTTAAGCAAAAATTTGCCCTTATTCAGGTGTGGGTACGCCACTTTAC	1093
O	1061	GGGCGCTACGACGAGTACGCTCCCTTTAGCGGCTCGACGAAGCGCTCGAGACGGCCCAT	1120
D	1094	CAGCTTATTTACACACATCACTCTTTATTAATGCCATTTGTGAGAAATTAATCTTATTTCCA	1155
O	1121	CGTACCCGTGGCGCTACACACGCTTCT-----C	1155
D	1154	AATTTGATATTAATCAATCAACTCAAAATATTAATTTGGTCCAAAAGCATTAATTAATTTAC	1213
O	1154	CCATTTACGGGACAGTGGCTCAACGCCGACGGGCGCTCGGGCATGGCGTGGAGGTC	1213
D	1214	CAGATTAATAGTCCACAAGATATCAATCCAAAAGTGTAAACCCGGGTTTCCATGAAT	1273
O	1214	TCAA-----CGAGCCCGCTGATACCCCTTAACCGCCAGACACATTTGACGAAGCTTCT	1264
D	1274	TTTATTAATAAAGCCAAATTCGGTTCTCTAATGAAATAAGAAATTAATTTGATGAATTAAGATA	1333
O	1285	TCACCAAGACGGACATCGACCTGTGTGACTACTACCAACCCCAAGGCGGACAGACAGTGG-	1322
D	1334	CAGAAATATCTGATATTTATTTAGGATTAATATATCATMAAGATATCCCTTTAATGGGT	1393
O	1334	-----TAGCGGCACATGAGAGGACGTACACGCCGACGAGGATCGACCTTCGAGTCG	1378
D	1394	TATATTTATTTGATTTGGAATTTGGAATTAACACCATCAAAAAACCAACATTATGAAATTTG	1433
O	1379	GCTCTGCTGTGCGGACGCGCAAAAGCGTACGTACAGCAGCCAGTCTGTCTGCAACAG	1438
D	1454	GTTTAACTGTTCAATGATACAGCACAATTTATTTATGATGATTAACCTTTGTGTTGATATA	1513
O	1439	CCACCAAGCAGGTCGCCGGGAGTCCCTGTGCGGCTCCGCCACCGGACAGAGACAGGGCC	1498
D	1514	AAACCAACAACAATTAAGAGGTTCTATTTTGTGAATTTGGAACCAATTGAAGAACAGGAT	1573
O	1499	GCATCAATCTGTCACAGGGACACACCTACAGTTCAAGTTCAGTTCCGCTCCGACCCA	1558
D	1574	CAATTTGAATTAATCAAGGATTAATTAATTAATCAATTTGTAATATGTTGGCGACCTA	1633
O	1559	CTTACACCTTCAAAGGGGACACCATGTGTCGCCGGCACGGCTCCCTCCGCGTGGCGGCT	1618

Db	1634	CATTTACCTTAAAGATCAAGTATGGAGAAATATTTTGGTGCGATTCGTTGTGATGA	1693
Oy	1619	GCAAGTCATTGAGACACAGGCC---GAATGGAAAAGTCGTGCCCCCTGCCACAGGAGC	1675
Db	1694	ATGAATATTTTAAAGATGAATGAACAGAAATATCATCAATTCACATAATTTTGGTCTAAATCAG	1753
Oy	1676	ACGACACAGTCATCATCTGCGCGGCCCTTAACGCCACGTACGGGAGACGAGGCGGCCGAC	1735
Db	1754	TAGATTATAGTATTTTATTTATTTATTTAGTATTAATAAAGATTGGGAAAGTAAATCTTATGATC	1813
Oy	1736	GCGGAGCATGAAGCTCCCGCGCGTGCCTGAGACACTATTTGCCGACGTGGCCGCCGCGA	1795
Db	1814	GTCGATATGAATAATTCAGTATTCAGAGATTAATATGATTAATGATGATGACGTTA	1873
Oy	1796	ACCCAAACCCGTCGTCATGCAGACGGGACACCCCGAGAGATGTC--CTGCTCG	1852
Db	1874	ATCCCAACACCATTTATTTGTAATCAATCCGATACACCTGTGGAATTTCTGAATGGTGG	1933
Oy	1853	ACGGCAGCCCGCGCTCATCCAGGCGCTGTACGGGGGCAAGAGACGGGCAACTCCATTG	1912
Db	1934	ATTAAGCTGAAGCTTTGGTTCATTTGGTATGATGGGGAATGAAGTGGATGCCATTG	1993
Oy	1913	CCGACGTCGTCCTTTGGGACATCAACCCCTCGGGAGCGTGCCTGACGTCCCAAGC	1972
Db	1994	CTGATATTTTATTTTGGTAATGTCAATCCTAATAGGGAATATCATTTGACTTTCCCAATGA	2053
Oy	1973	GCTGCAGAGCAACCCCGGCTTCTCAACTTCCGACCGAGGCGGGGCGCACGCTGTACG	2032
Db	2054	AAATATATGATTAATCCAACTTTTGAATTTTAAATGAAAGAGGTCAGATATTATATG	2113
Oy	2033	GGGAGGACGCTACGTCGGGTACAGGTACTACGAGTTTGGCCGACAGGAGCTCAATTTCC	2092
Db	2114	GAGAAAGATATTTTGTGTGTATAGATTAATTAAGAAAATTAAGAAATGCAATGTCAGATGCTTTC	2173
Oy	2093	CCTTTGGCAGCGCTGCTCATACACACTTTTGGCTTTTCCAACT	2138
Db	2174	CATTGTGATTTGGATATCATATACAAATTTGAATTTTTTTGATTT	2219

RESULT	12
LOCUS	SC8A6
DEFINITION	Streptomyces coelicolor cosmid 8A6.
ACCESSION	AL031013 AL645882
VERSION	AL031013.2 GI:20520810
KEYWORDS	beta-glucosidase; elongation factor G; fuaA; succ; succinyl-coa synthetase alpha chain; succinyl-coa synthetase beta chain; succ; rpp-requiring enzyme; transfer-RNA-Leu; TTA Leu codon.
SOURCE ORGANISM	Streptomyces coelicolor A3(2).
REFERENCE AUTHORS	Streptomyces coelicolor A3(2) Bacteria; Firmicutes; Actinobacterias; Actinomycetales; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces. 1 (bases 1 to 37445) Redebach,M., Kiser,H.M., Denapate,D., Eichner,A., Cullum,J., Kinashi,H. and Hopwood,D.A. A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome Mol. Microbiol. 21 (1), 77-96 (1996)
JOURNAL MEDLINE	97000351
PUBMED	8843436
REFERENCE	2 (bases 1 to 37445) Seeger,K.J. and Harris,D.
AUTHORS	Unpublished
JOURNAL	3 (bases 1 to 37445)
REFERENCE	Parkhill,J., Barrell,B.G. and Rastandream,M.A.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUL-1998) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood. [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT	On May 9, 2002 this sequence version replaced gi:3288600.
COMMENT	Notes:

Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.

Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

(URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>) CDS are numbered using the following system eg SC7B7.01c, SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/>

Jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (atc)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 8A6 lies between 3F9 and 1F2 on the AseI-A genomic restriction fragment.

FEATURES

location/Qualifiers

1. 37445

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid 8A6"

misc_feature

2. 107

/note="nominal overlap with S8A6 from 1 to 106

/note="nominal overlap with cosmid 3F9 from 19725 to 19830"

657. 711

/note="hairpin loop with 23 bp stem"

/complement(730. 2160)

/gene="SCO6581"

/complement(730. 2160)

/gene="SCO6581"

/note="SC8A6.02c, probable transmembrane transport protein, len: 476 aa; some similarity to e.g. TR:O51330 (EMBL:U4007 5) Oxalate:formate antiport protein from

Oxalobacter formigines (418 aa), fasta scores: opt: 226

z-score: 230.1 E(): 1.4e-05, 24.1% identity in 436 aa

overlap"

/codon_start=1

/transl_table=11

/product="putative transmembrane transport protein"

/protein_id="CA19774.1"

/db_xref="GI:3288602"

/db_xref="SPTREMBL:087836"

/translation="MGQALARSFADGAIADMTTDTVTRAAVREYTDNRGRYRVE

SDIDLMGRKRWVITLPMIGMGISSAEVAFASAEPLLEAQMSSGSIYMTAAVAF

COAAVAPARLRRESCKLPARMMMLISAGTLLGYSLEAPAPVSLAFTIGFVGSNG

AGAVYATCVNMGKWPYERGRGKTFVNGFAGVSEVFEVIFGMDTSMFRVAVLSA

GVPLAAVAVATAGFPRDPKNNMPAAVDPLNPADPARSLKNNPAPVAVSPVMEW

KTRGVALLMRECLACTSGVNFETAFQVYDIEEAGFAPAAASIKATVNGRGVY

GWSDLDLGRKQCLLYVCAITLLGAFGLTMSAEIKNLPLVFSATISFGGCAFPFPA

ALTADYFGENNNAITNGVYSKLVSLGLAGMGSVYGVNGAFSLAGSISIFPAGF

VALSLRPPGRKRVTPNPDLGEG"

/complement(2233. 4377)

/gene="SCO6582"

/complement(2233. 4377)

/gene="SCO6582"

/note="SC8A6.03c, unknown, len: 714 aa; similar to many

hypothetical proteins including several from each of the

misc_feature

gene

CDS

/note="PS00216 Sugar transport proteins signature 1"

/complement(4382. 5614)

/gene="SCO6583"

/complement(4382. 5614)

/gene="SCO6583"

/note="SC8A6.04c, probable transferase, len: 410 aa; highly similar to TR:006644 (EMBL:U82167) formyl-CoA

transferase from Oxalobacter formigenes (428 aa), fasta

scores: opt: 95.4 z-score: 1732.0 E(): 0.50.7% identity

in 428 aa overlap. Also similar to many dehydratases e.g.

CAIB_ECOLI L-carnitine dehydratase (EC 4.2.1.89) (405

aa), fasta scores: opt: 285 z-score: 317.7 E(): 1.9e-10,

26.2% identity in 427 aa overlap"

/codon_start=1

/transl_table=11

/product="putative transferase"

/protein_id="CA19776.1"

/db_xref="GI:3288604"

/db_xref="SPTREMBL:087838"

/translation="MTAKALEGIRVLDMTNHOQSPATQLTAMIGADYVLEAPGDI

TRGQRLDLPDVSLEFMTNCKRSITNTKSEKKEITLIRSDVAVENVEFGAV

DMGQRLDLPDVSLEFMTNCKRSITNTKSEKKEITLIRSDVAVENVEFGAV

ATGAOIGDSCTGVHYVAGTALAAYREHGRGQARAPGQENDYIYIVPQNPISL

HQPLAEYFNEDEYPRKSNAGSGGQPMAYKACAPVGNQYIYIVPQNPISL

IQRLAEYFNEDEYPRKSNAGSGGQPMAYKACAPVGNQYIYIVPQNPISL

KEIIDDSLVANEMVTVVPHPERGEVTSPLKIDSVEVTSPLGHEHNEVYVG

ELGIDDEELRLKSSGVIT"

/complement(5623. 5627)

/note="possible RBS upstream of SC8A6.04c"

/complement(5626. 7308)

/gene="SCO6584"

/complement(5626. 7308)

/gene="SCO6584"

/note="SC8A6.05c, probable TPP-regulating enzyme, len: 560

a; similar to many e.g. IIVB_MYCU probable acetylacate

synthase (EC 4.1.3.18) (547 aa), fasta scores: opt: 733

z-score: 1305.6 E(): 0.36.4% identity in 557 aa overlap.

Contai ns P500187 Thiamine pyrophosphate enzymes signature

and Pfam match to entry PF00205 TPP enzymes, Thiamine

pyrophosphate enzymes, score 343.90, E-value 5e-106"

/codon_start=1

/transl_table=11

arche ons Archaeoglobus fulgidus TR:028341 (EMBL:AE000970)

conserved hypothetical protein AF1938(673 aa), fasta

scores: opt: 960 z-score: 1289.8 E(): 0.35.6% identity in

693 aa overl ap and Pyrococcus horikoshii D1028624

(EMBL:AB009518) hypot hetical protein PHV016 (460 aa),

fasta scores: opt: 1280 z-score: 1246.9 E(): 0.44.38

identity in 461 aa overlap. Co ntains P500216 Sugar

transport proteins signature 1. Also s imilar to S.

coelicolor hypothetical protein SC9B10.09 (E(): 6.1e-15,

33.9% identity in 375 aa overlap)"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="CA19775.1"

/db_xref="GI:3288603"

/db_xref="SPTREMBL:087837"

/translation="MAEPRLARSGILDVAGRGRTALTAPGKVLADAVIAYGEE

LARDVDEAVACARRGVYVVKIYSPDLIHTDAGGVYVGEAGVADPRAFCRVNDA

RAYDASARIIEVOJLOELPQGOEVIVGAVTDPFGKVVAFGLGVLEVDVTFRLA

PVDADALSMDSIRAAEVLKGVNGQAGVDMVAEOLIRVSELVADPELAEVDLNP

VIAPEGAIVADIRVILATDTPKERRRYREELTSMRLMOPASVAVTIGASGPGKI

GSVNRNIVDGGEGEIHVPNKADDIIGRAKYSVTDVPEGVAVVATAPRAVAA

LEEVGRKRIPANVILIPSGFAETGBOALODEIVAAERHGVLELGNITGYISTMDDC

AFECTPYVKGCVALTSSGSGIGMAITFCFARTTGTGSAITGLCNKSLDDDDILTFE

GEDPTECIAMHLEDLKDGRAVEEAKATVPKPYVLAAGRTAGAKAKASHGALA

GDADVYDILRQAGVIRAPAGLNEMLEKRALPLPAPGDNVVIITGAGSGVILSDA

VTDNLSTMEIIPDIDRAFAFIPFGAGALPVDITGEPSTYEATIRLDEDPRIH

ALVIGVHTIYTPPVAFELTARVAAEEREGVEPVVASTAGVEVEACQYLFERG

VVAVYETKEKPVAVIGAVKRAAAGLLGGGS"

/complement(3340. 3387)

/gene="SCO6582"

/note="PS00216 Sugar transport proteins signature 1"

/complement(4382. 5614)

/gene="SCO6583"

/complement(4382. 5614)

/gene="SCO6583"

/note="SC8A6.04c, probable transferase, len: 410 aa;

highly similar to TR:006644 (EMBL:U82167) formyl-CoA

transferase from Oxalobacter formigenes (428 aa), fasta

scores: opt: 95.4 z-score: 1732.0 E(): 0.50.7% identity

in 428 aa overlap. Also similar to many dehydratases e.g.

CAIB_ECOLI L-carnitine dehydratase (EC 4.2.1.89) (405

aa), fasta scores: opt: 285 z-score: 317.7 E(): 1.9e-10,

26.2% identity in 427 aa overlap"

/codon_start=1

/transl_table=11

/product="putative transferase"

/protein_id="CA19776.1"

/db_xref="GI:3288604"

/db_xref="SPTREMBL:087838"

/translation="MTAKALEGIRVLDMTNHOQSPATQLTAMIGADYVLEAPGDI

TRGQRLDLPDVSLEFMTNCKRSITNTKSEKKEITLIRSDVAVENVEFGAV

DMGQRLDLPDVSLEFMTNCKRSITNTKSEKKEITLIRSDVAVENVEFGAV

ATGAOIGDSCTGVHYVAGTALAAYREHGRGQARAPGQENDYIYIVPQNPISL

HQPLAEYFNEDEYPRKSNAGSGGQPMAYKACAPVGNQYIYIVPQNPISL

IQRLAEYFNEDEYPRKSNAGSGGQPMAYKACAPVGNQYIYIVPQNPISL

KEIIDDSLVANEMVTVVPHPERGEVTSPLKIDSVEVTSPLGHEHNEVYVG

ELGIDDEELRLKSSGVIT"

/complement(5623. 5627)

/note="possible RBS upstream of SC8A6.04c"

/complement(5626. 7308)

/gene="SCO6584"

/complement(5626. 7308)

/gene="SCO6584"

/note="SC8A6.05c, probable TPP-regulating enzyme, len: 560

a; similar to many e.g. IIVB_MYCU probable acetylacate

synthase (EC 4.1.3.18) (547 aa), fasta scores: opt: 733

z-score: 1305.6 E(): 0.36.4% identity in 557 aa overlap.

Contai ns P500187 Thiamine pyrophosphate enzymes signature

and Pfam match to entry PF00205 TPP enzymes, Thiamine

pyrophosphate enzymes, score 343.90, E-value 5e-106"

/codon_start=1

/transl_table=11

[illegible]

DEFINITION	Streptomyces narbonensis desosamine biosynthetic gene cluster, partial sequence.
ACCESSION	AF521878
VERSION	AF521878.1
KEYWORDS	GI:22001230
SOURCE	
ORGANISM	Streptomyces narbonensis. Streptomyces narbonensis Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE	1 (bases 1 to 17665) Bate,N. and Cundliffe,E. The desosamine biosynthetic cluster of Streptomyces narbonensis, producer of narbomycin
AUTHORS	Unpublished
TITLE	2 (bases 1 to 17665) Bate,N. and Cundliffe,E.
JOURNAL	Submitted (14-JUN-2002) Biochemistry, Leicester University, University Road, Leicester LE1 7RH, UK
REFERENCE	Direct Submission
AUTHORS	
TITLE	
JOURNAL	
FEATURES	Location/Qualifiers
source	1..17665
gene	/organism="Streptomyces narbonensis" /db_xref="taxon:67333" <1..528
CDS	/gene="nbma" <1..528 /gene="nbma" /note="terminal thioesterase I (TEI) domain of type I polyketide synthase" /codon_start=1 /transl_table=1 /product="Nbma" /protein_id="AAM88350.1" /db_xref="GI:22001231"
gene	/translation="AVLRAAGDAPVVLHGSGALLAHGLALERLERAGTAPAGIVLYLDPYAPGHDEPIEWSROLGELFAGLEPELSELRLLAMGRVAFPLAGPPGSSAPVLYRAERLGGMLPEERGGMRAMHMDLPHSVADVPCDHTTMRDRAPAVAVLAMDGITRDTAGRTAGAGAQ"
CDS	525..1370 /gene="nbmb" 525..1370 /gene="nbmb" /note="thioesterase II (TEII)" /codon_start=1 /transl_table=1 /product="Nbmb" /protein_id="AAM88351.1" /db_xref="GI:22001232"
gene	/translation="MTRVLVNDSSILMIRRHPSPNSAVRLYCLPHAGSASYPFRFEELEHPSEVALSVQYPGRODRAPCESEVELAEHVVAATEAMREGLAFEGHSIGDENSVAYEARLIEORHGVREGLYVSGRRAPSLAPDRLYHQLDRAFIARLROLSGDERFIQDDELRLRYLYPALRSDYKAATFYEHRSKALACPVAIACDRPKAPLVEVEMKRWTSPEFLCRATSGGHFYLANDQMHCINDISDHLVTRGAVDSRVVQPTRLIEGAARKWQNR"
CDS	1367..2713 /gene="nbmc" 1367..2713 /gene="nbmc" /note="putative NDP-deoxyhexose 3,4-isomerase" /codon_start=1 /transl_table=1 /product="Nbmc" /protein_id="AAM88352.1" /db_xref="GI:22001233"
gene	/translation="MTDDPARALTOPPRGLRTYSANADPELCAHLLLETRGITHANGNDPIAAVIRGQADDPYATERVARRGVLSPTGWSVNTADHVAASVLCSTEEFGPGADGVAPQDVQLSYGEGCPLEHEDQAPRAAAEYVGAQORAAVEEYHRLTEGLAFDPGADSYAELELGFVPAVGAAGAAVLYGVAGRGFEALYERLPLSDLSLAPSLATPVAAGDALLATLLAEAGVSAAATGSDPSGSPSSDPSASSGPTASGGTATPSSGVAGRGDGLALSALGTAAYOULTGNNAVIALLAHEQORETCARPGIAAAVETTLAYDDPVOLDARARVETELACRRRLPAGAHVYVLTAGIGRDGVFTDPERFDLARPATAHLLHPHAGPVSIVLRLOAEVALRTLAGRPFGKRAQVLRPRAPVGRGRLSPVGS"
gene	2782..4062

```

CDS      . /gene="nbmd"
          2782..4062
          /gene="nbmd"
          /note="desosaminyl transferase"
          /codon_start=1
          /transl_table=11
          /product="Nbmd"
          /protein_id="AA08353.1"
          /db_xref="GI:22001234"
          /translation="MRVLTSPAHHTHYGVLPLAMALLAGHEVRVASOPALDTIT
          GSGLAAPVGTDLIHEYKVRWAGEPRNHPALADEKRSPLDMDHALGIDATLARY
          FYLIANDSMIDLVGFARSWQPDVLMEPTTYGAAVAQYTGAAHARVLMGPVMS
          ARKFAALDRQPEHREDPTAEMLTWLTLDRGASFESEELTGQFVDPTEPSRLDT
          GLPTGVGVVYVNGSVYDMLAEPPARVCLITGVASREVLGDSQADILLEALA
          DDIDIEVALIDAGORAEIRSPKHTRTDFVPMHLLPSCSAIVHGGAGTYATAVIN
          AVQVLAELMDAPKARAVAROGAGCPPLPAGLPGQAVRDAVVRILDDPSVLAAGR
          LRETHRDPTPAGIVPELERLAOHRPPADARR"
          4151..4864
          /gene="nbme"
          4151..4864
          /gene="nbme"
          /note="N,N-methyltransferase"
          /codon_start=1
          /transl_table=11
          /product="Nbme"
          /protein_id="AA08354.1"
          /db_xref="GI:22001235"
          /translation="MYEVDHADVDFLYLGKRDYAAEASDIADLVKRTTPAESLLD
          VACGCTHLEHPTKEFGDPALELSEDMTHARKRLPDATLHQGDMRDFRLGRSAV
          VSMFSVGYLRTTAELDAAYASRAHLEPGVYVPEPPEFADGWSAVYVRDG
          RTVAVSHSVRDEGATRMVEHFTVADPGKGVRFSDVHLITLFHAEVEAFTAGLR
          VEYLEGSGSGRLFVQVPA"
          4908..7424
          /gene="nbmf"
          4908..7424
          /gene="nbmf"
          /note="beta glucosidase"
          /codon_start=1
          /transl_table=11
          /product="Nbmf"
          /protein_id="AA08355.1"
          /db_xref="GI:22001236"
          /translation="MHOAREKSTVTGKTRIPRVRSRTTPRAFTLAVVGTLAGTIV
          AAAPGAGCTGHVOTYSKAEIVAOQMTLDEKISFVHMLDPPRONGVYLPGRLEIP
          BLRADGNGRLVGRTRATAPAYALASFPDMSADSVGMGRDGRALGODMVGPP
          MMNRIYVPHGRNTEFSEDLVSSRAVAQITGICGAGIMTTAKHFAANNOONNES
          VNAIVDEQTLREIEFPAEASKAGASFWCAINGVNGKSCGNDELNNVLETKWGF
          OGWMSDMLATPGTDAITKGLDOMAGVELPGDIPGEPSPAKFEGDALKQAVLNGTY
          PEAAVTRSAERIVNOMDFGLLATPAPRERDPAQAQAVSRKVAENGAVLLNBSGA
          LPIADGAKSTAVIGPTAVDPKVTGLSAHVPPSAAAPLDTTKARAGAVTYETG
          BETPSTRPAQSLVIGPNOGHOLEPGKAGALYDGLTVPADGEYRISVKAAGATVYEG
          LGSHTIEGQYKYSPLKLTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK
          KAVESARKARAIYVADDDGEGVDRPLSLPGTODKLISAVDANANTIVLNTGSS
          VLMPSIKTRAVLDMWYPGOGAETATALLIGDYNPSGKLQSPRAENONHAYAGDPN
          RYPGVNOQYSEGIHVGYSRWFDEKNKPLPEFGHGLSYSTQSATFYVTRTGGICK
          VYTVVNSGORAGCEVQVAYLGASPKYAPAEKKLVGYKVALAAGESKTVTVNDR
          ROLQYWDASDSMRTGSRLLIOTGSSADLKSSQYVW"
          complement(7514..8653)
          /gene="nbmg"
          complement(7514..8653)
          /gene="nbmg"
          /note="aminotransferase"
          /codon_start=1
          /transl_table=11
          /product="Nbmg"
          /protein_id="AA08356.1"
          /db_xref="GI:22001237"
          /translation="MTARTETGRVPLDLKAAVEELREPTDAIARVDSGRYLLGE
          LEAFEAFAFYCECTDAHGVNSGMDALDLALGLTHIGDEYIYPSHYTASMLAVSA
          LGATPVPVENEDHPTLDPLVEKAITPRTALDLVIGHADIDALRDLADRGHLH
          TVDDAQAAGARRGRRIGAGASVAASFYPGKNGSCFGDGAAGVVTGPALAEIRML
          RYNGSROKYSHEKTKNSRLDEIOAAVLEIRAHNDSDNNGRSTLAAYISGLGLSP
          IALPVTAPGIDPVMHLFTVTRIDRDELROHDSRGVDLTITHYVPVHLSPAYAGAPP

```

```

gene      EGSIPRAESFAROYLSLPIGPHLERTQALRVIDAVERMAQAGR"
CDS      complement(8650..9684)
          /gene="nbmd"
          complement(8650..9684)
          /gene="nbmd"
          /note="NDP-glucose 4,6 dehydratase"
          /codon_start=1
          /transl_table=11
          /product="Nbmd"
          /protein_id="AA08357.1"
          /db_xref="GI:22001238"
          /translation="MRLLVTGAGCTGSHFYROLLAAYPDVAGEVYILDSLYAGN
          RANLAPVEADPRLREHGDIRPADLAELEKGVDAVHFAESHYDSSIAGASVFTGT
          INOQOTLQCLADGAGVAVVSTDEYGSIDGSVMSSPLEPNSPTASKRAGSDI
          VVARHRYRGLVTRITROCNNGPYQHEEKLIPLEVMILDGGLPLYGIDANREW
          HHDHCRGAIYALAGRAGEYHIGGRELINRELITGLIDSLSDAMS SVRARIDRG
          HDIRTSLDGKTIKRELGTPQIAFPADGLARTIRWRRENRDMEPLKATAPOPPATADQ
          PPATAVGSA"
          complement(9759..10637)
          /gene="nbmi"
          complement(9759..10637)
          /gene="nbmi"
          /note="NDP-glucose synthase"
          /codon_start=1
          /transl_table=11
          /product="Nbmi"
          /protein_id="AA08358.1"
          /db_xref="GI:22001239"
          /translation="MKGIYVLAGSGSTRPLHPTATSVISKIILPYNKPMTIYPLSLMLG
          GIREIOTISPHOIELFOSILNGRHLGIELDYAVQKEPAGIADLILVAGHIGDGT
          ALIDGDIHFGRLYSLSDLSIRLKGCVLFGVPYRDEPRGVAEVDATGRLTGVEK
          PERPSRLAVATGLITDNDYDIAKNIIRSPGRELITPVNRYLERGAEALVDLGRG
          FAWLDGTSDLSLRAAQYVAVLEERGVMIAGLEIAPFMGTIDAEACHSLGEGLSRT
          EYSGYIMEIARGEGAP"
          complement(10840..12297)
          /gene="nbmj"
          complement(10840..12297)
          /gene="nbmj"
          /note="putative NDP-deoxyhexose 3,4-reductase"
          /codon_start=1
          /transl_table=11
          /product="Nbmj"
          /protein_id="AA08359.1"
          /db_xref="GI:22001240"
          /translation="MTAPAPVATPSPARCAPHGADLGAARAVAGOSIAAGLVPDSDEA
          GSTAHHLGLVRRGNSEPTPLPEARRDLDVDRAPRRLALPPEVELAAVAGGSA
          GAYWONTLLPLERGVDSALAGPVPYVSYGLTPPGIGTLAAHAAHEGLRBPVYTNSE
          LDGTNAMERSYIDEIPAGNPAITFSGLEPLTPGIGTLAAHAAHEGLRBPVYTNSE
          ALTEETLERQPLGLHAHVRISLGLNDEYEDTGTAKAFRKYRELRFPQOLRAR
          EAPILGFRAYIVLPGRAHRLIDVDFIADIDESQGRITDVNIREDYSEREGKLPQ
          EERATLOALAAFERVEREPTPGHIDYGVALNSLKAGAPAEILRIKPAVMPRTAHPQ
          VAVOVDLIGDVIYIRENGFEPDLCATRIAGRVGPGSLTEVVADVVERGVRVARDG
          DEYFMDGFDQVYATRLNLENDIDMGHEARGLR"
          7.4%; Score 184.4; DB 1; Length 17665;
          Query Match 48.2%; Pred. No. 1.4e-13;
          Best Local Similarity 956; Mismatches 15; Gaps 12;
          Matches 1028; Conservative 0; Indels 150; Gaps 12;

```


TITLE Analysis of eryB1, eryB11 and eryB17 from the erythromycin biosynthetic gene cluster in *Saccharopolyspora erythraea*

JOURNAL Mol. Gen. Genet. 258 (1-2), 78-88 (1998)

MEDLINE 98273631

PUBMED 9613575

REFERENCE 2 (bases 1 to 2760)

AUTHORS Leadlay, P.F.

TITLE Direct Submission

REFERENCE Submitted (25-JUL-1997) P.F. Leadlay, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QM, UK

REMARK Updated by [3]

AUTHORS 3 (bases 1 to 2760)

TITLE Leadlay, P.F.

REFERENCE Direct Submission

REMARK Submitted (18-MAR-1998) P.F. Leadlay, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QM, UK

COMMENT On Mar 25, 1998 this sequence version replaced gi:2765397.

FEATURES

location/Qualifiers

1..2760

/organism="Saccharopolyspora erythraea"

/strain="NRRL2338"

/db_xref="taxon:1836"

1..45

/gene="ermE"

/gene="ermE"

<1..45

/gene="ermE"

/codon_start=1

/transl_table=1

/protein_id="CA474701.1"

/db_xref="GI:2765398"

/translation="TGGPGRGGPGOR"

/complement(100..2526)

/gene="eryB1"

/complement(100..2526)

/gene="eryB1"

/codon_start=1

/transl_table=1

/product="beta-glucosidase"

/protein_id="CA474702.1"

/db_xref="GI:2982611"

/db_xref="SPTRMBL:070021"

/translation="WTGGERVRLVIRLAPLLVPLVAVSPVRSQVDELIGOL
TIDELTSFYVDYNEKDIPLAKLIRGPRIGIPIDRDPAGVTHIOPAIAPAPVA
TASAFDRLAHETGVLCREARFEDDITIGPMNIRVOAGNEFTFSEDELAVAR
TAAQIRGHSQMTSAKHYAANTOEIDRTIDVDIDRLELRELEAEAVAGA
TSVCAVPRKVTAGCGRHRLTELEKQWCFKGVSDMTATHTATDIAAGIDEMG
VEVRDGSIFRKYLGELAKIRREGRIPEALDASVRIITQFERGLDETRPRP
ERDVAGTRIAOEVAESGAVLRNKGVLPLDPAAGDIAVIGSAQPPVTLGSGSY
VPPDANAPLDTITQVSGSGVGVSEELKGAIPETALQPAFVGVLTGKEHTI
IYDGRUKADADLYRIARIDGNSLIDGAPLVGDFGPLTSVPLVLTGKEHTI
OMTGAPEVGGSLVDYDLTVTPGHAQREFDAVRARDSDVAVFAVDDGAEATTA
LSPGIDKLIDAVASVNPNTVYLTGSSVTMRLDKTRAVLDMVPGGAATTA
LLEGAPEGRGLTQEPVSOERTVPGVPARPGVDGKVIHSEITFGSYRMYDEGVL
PLFEGHGLSTYFERTDPEVETRDGLDVTYVANTQKRGSDVVOYGLPSPQVYL
DOAPRLAGYCKVELAPGETKRVHVAERALQHWDEAGMKLGKRAVEIGSSSR
DIDLADINL"

2716..2760

/gene="eryB11"

/gene="eryB11"

2716..2760

/gene="eryB11"

/codon_start=1

/transl_table=1

/protein_id="CA474703.1"

/db_xref="GI:2765400"

/db_xref="SPTRMBL:000463"

/translation="MTFLVGLGCRICGN"

BASE COUNT 383 a 966 c 982 g 429 t

ORIGIN

Query Match 7.1%; Score 178.2; DB 1; Length 2760;

Best Local Similarity 54.0%; Pred. No. 1.2e-12;

Matches 416; Conservative 0; Mismatches 343; Indels 12; Gaps 2;

QY 1591 GGCACGAGCTCTCTCCGCGTCCGCGGCGGCTGCAAGGTCATGACGACGAGCCGAATCGAA 1650

DB 978 GGGCGGGTGTGCTGCACAGCTGCAGCTGACCTGAGGTACCCGGGGAGACCCAGCGCAG 919

QY 1651 AAGTCCGTGCGCTCCGCAAGAGACGACGACGATCATCTGCGCGGCGCTTAACCC 1710

DB 918 TTGCACGCGCGGTGGAGGCGGCGGCGGACCTCCGACGTCGCGGTGGTTTTTCCCTACAG 859

QY 1711 GACTGGAGACCGAGGCGCGGACCGCGGAGCATGACCTCCGCGGCTGCGACAG 1770

DB 858 GACGCGCGG---GAGACCGCGGACCGGACGTCGCTGCTGCTGCAATGACAGCGGAC 802

QY 1771 CTCATTGCGGACGCGGCGCGGCGGACCCAAACCGCTGCTGCTGCAATGACAGCGGAC 1830

DB 801 CTGATCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742

QY 1831 CCGGAGGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1890

DB 741 TCGGTACAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682

QY 1891 AACGAGACGCGGCACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950

DB 681 CAGGCGGGGCGAGAGGCTCACCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 622

QY 1951 CTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2001

DB 621 CTCACCCAGACCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562

QY 2002 TTCGCGACCGAGCGCGGCGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2061

DB 561 TTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502

QY 2062 TAGAGATTGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2121

DB 501 TACGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 442

QY 2122 TTTGCTTTTCCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2181

DB 441 TTGAGACGCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382

QY 2182 GTGAAAGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2241

DB 381 GTGCGAAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322

QY 2242 CAAGCGGCAAGATTACCGCCCGCTCAAGAGGCTCAAGGCTTGGCAAGGCTCAAGTGC 2301

DB 321 CCGCAGGTTCCGCTGAGCAGGACCGCGGCGGAGCTGCGGGTACCAAGAGTGCAGACTC 262

QY 2302 CAGCCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2352

DB 261 GCAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 211

Search completed: May 4, 2003, 03:05:15

Job time: 4797.77 secs

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry, no matter how small, should be carefully documented to ensure the integrity of the financial data. This includes recording dates, amounts, and the nature of the transactions.

The second part of the document outlines the procedures for reconciling the accounts. It states that a thorough reconciliation should be performed at the end of each month to identify any discrepancies between the recorded transactions and the actual bank statements. Any differences should be investigated and corrected immediately.

The third part of the document provides a detailed breakdown of the expenses incurred during the period. It lists various categories such as salaries, rent, utilities, and supplies, and provides a clear summary of the total costs for each category. This information is crucial for understanding the overall financial performance of the organization.

The final part of the document concludes with a statement of the total net income or loss for the period. It reiterates the importance of transparency and accountability in financial reporting, and expresses confidence in the accuracy of the presented data.